

Language, culture & cognition

The Anatomy of Meaning

Speech, gesture, and composite utterances

N. J. Enfield



This page intentionally left blank

The Anatomy of Meaning

How do we understand what others are trying to say? The answer cannot be found in language alone. Words are linked to hand gestures and other visible phenomena to create unified 'composite utterances'. In this book N.J. Enfield presents original case studies of speech-with-gesture based on fieldwork carried out with speakers of Lao (a language of Southeast Asia). He examines pointing gestures (including lip and finger-pointing) and illustrative gestures (examples include depicting fish traps and tracing kinship relations). His detailed analyses focus on the semiotic unification problem, that is, how to make a single interpretation when multiple signs occur together. Enfield's arguments have implications for all branches of science with a stake in meaning and its place in human social life. The book will appeal to all researchers interested in the study of meaning, including linguists, anthropologists, and psychologists.

N.J. ENFIELD is a senior staff scientist in the Language and Cognition Group at the Max Planck Institute for Psycholinguistics, The Netherlands. His recent publications include *Person reference in interaction* (with T. Stivers, 2007, Cambridge), *Roots of human sociality* (with S. C. Levinson, 2006) and *Linguistic epidemiology* (2003).

Language, culture and cognition

Editor

Stephen C. Levinson

Max Planck Institute for Psycholinguistics, Nijmegen

This series looks at the role of language in human cognition – language in both its universal, psychological aspects and its variable, cultural aspects. Studies focus on the relation between semantic and conceptual categories and processes, especially as these are illuminated by cross-linguistic and cross-cultural studies, the study of language acquisition and conceptual development, and the study of the relation of speech production and comprehension to other kinds of behaviour in cultural context. Books come principally though not exclusively, from research associated with the Max Planck Institute for Psycholinguistics in Nijmegen, and in particular the Language and Cognition Group.

- 1 Jan Nuyts and Eric Pederson (eds.) *Language and Conceptualization*
- 2 David McNeill (ed.) *Language and Gesture*
- 3 Melissa Bowerman and Stephen C. Levinson (eds.) *Language Acquisition and Conceptual Development*
- 4 Gunter Senft (ed.) *Systems of Nominal Classification*
- 5 Stephen C. Levinson *Space in Language and Cognition*
- 6 Stephen C. Levinson and David Wilkins (eds.) *Grammars of Space*
- 7 N.J. Enfield and Tanya Stivers (eds.) *Person Reference in Interaction: Linguistic, cultural and social perspectives*
- 8 N.J. Enfield *The Anatomy of Meaning: Speech, gesture, and composite utterances*

The Anatomy of Meaning

Speech, Gesture, and Composite Utterances

N.J. Enfield

Max Planck Institute for Psycholinguistics, Nijmegen



CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS

Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, São Paulo

Cambridge University Press

The Edinburgh Building, Cambridge CB2 8RU, UK

Published in the United States of America by Cambridge University Press, New York

www.cambridge.org

Information on this title: www.cambridge.org/9780521880640

© N.J. Enfield 2009

This publication is in copyright. Subject to statutory exception and to the provision of relevant collective licensing agreements, no reproduction of any part may take place without the written permission of Cambridge University Press.

First published in print format 2009

ISBN-13 978-0-511-51808-9 eBook (NetLibrary)

ISBN-13 978-0-521-88064-0 hardback

Cambridge University Press has no responsibility for the persistence or accuracy of urls for external or third-party internet websites referred to in this publication, and does not guarantee that any content on such websites is, or will remain, accurate or appropriate.

It is bad when one thing becomes two.

Yamamoto Tsunetomo, 1716

For Sam and Matt, and their worlds.

Contents

<i>Preface</i>	<i>page</i> ix
<i>Acknowledgements</i>	xi
1 Composite utterances	1
Part I Deictic components of moves	23
2 Demonstratives	25
3 Lip-pointing	68
4 Hand-pointing	90
Part II Illustrative components of moves	111
5 Modelling	113
6 Diagramming	149
7 Editing	194
8 Conclusion	221
<i>Bibliography</i>	229
<i>Index</i>	247

Preface

This is a book about meaning which concentrates mostly on the interpretation of speech-with-gesture composites. The payoff of studying speech-with-gesture is not only to understand gesture as a phenomenon of interest in itself, but – of greater consequence – to help shake linguists and other students of meaning from the view that language is an encapsulated system for conveying meaning. It isn't. What I've learnt from carrying out the work of this book is that meaning is composite and context-grounded no matter how you look at it. The type-level meanings traditionally described in linguistic semantics are never the full meanings that token utterances are taken by interpreters to have, or designed by producers to have.

Standing on the shoulders of a giant pyramid of other midgets, I offer a mere increment on the pioneering efforts of predecessors. The idea of composite utterances as promoted in this book is grounded in a significant prior literature. Those most proximally responsible for the perspective taken here include well-knowns like Adam Kendon, David McNeill, Herb Clark, and Chuck Goodwin, but also some less widely published and cited authors who have, nevertheless, had a direct effect on how I have come to approach the speech-with-gesture problem. Of particular note are Tatiana Slama-Czacu (who suggested that language and gesture could combine in a 'mixed syntax'), Randi Engle (whose neo-Peircean account of 'composite signals' showed how utterance meanings are not merely summations of their multimodal parts), and Arika Okrent (who insisted that if we are to understand how 'language' and 'gesture' differ and interact, we must define them in terms of semiotic function and not in terms of modality).

This book is one of very many contributions to an emerging discipline of research on speech-with-gesture. With no shadow of a doubt, the most urgent work for this new tradition is extensive and intensive description, across languages and cultures, and across types of social and communicative activity. Human gestures show extraordinary variation in form and meaning compared to those of even our closest relatives in the animal world. While it is apparently possible to summarize in a book chapter the repertoire of gestures of a single species of great ape (Call and Tomasello 2007), this is not

possible for our species. The diversity of social behaviour across human groups appears to be as great as that **between species** elsewhere in nature (cf. Dunbar 1988). So, we cannot confidently generalize about forms and functions of gestures and other meaningful bodily movements in human social behaviour until we have an idea of the degree to which these phenomena are variable across human groups. For this we need a deep descriptive tradition. Even if we decide that gestures are non-linguistic, they are nevertheless locally conventionalized and therefore widely variable. At present, too little is known. A major task, then, for gesture studies is to build an adequate descriptive base, as was (and continues to be) necessary for language in linguistic science. The greatest breakthroughs in our understanding of the universal properties of language came only after we had access to an extensive historical tradition of descriptive work on languages, large and small, from around the world. This history is yet to be written for gesture.

I dedicate this book to my sister and brother and the people and places who complete them. Sam and Matt are both one of a kind, and each, uniquely, half me.

N.J.E.

NIJMEGEN

Acknowledgements

The work presented in this book has arisen from several years of collaborative research on language, gesture, and cognition conducted together with members of the Language and Cognition Group at the Max Planck Institute for Psycholinguistics, Nijmegen. This work has been carried out within three projects of the institute: The Space Project, the Gesture Project, and the Multimodal Interaction Project. I thank my colleagues in each of these groups for their engagement, support, trust, and friendship.

For technical and research support, I thank the Technical Group of the MPI Nijmegen, for assistance with field equipment and for providing the means to conduct the kind of research reported on here. I am also grateful to student assistants in the Language and Cognition Group for their help with practical problems in the revision and formatting of the text and figures. In order of appearance: Norah Carp, Alex Dukers, Ludy Cilissen, Winie van den Bosch, and Annelies van Wijngaarden. I gratefully acknowledge the support of the Max Planck Society for their funding of this research.

Chapters 2–7 are significantly revised versions of previously published articles. I am grateful to the publishers of these articles for kindly giving me permission to revise and reproduce them here: Demonstratives in space and interaction: Data from Lao speakers and implications for semantic analysis, *Language* 79.1, 2003, 82–117 (Chapter 2); ‘Lip-pointing’ – a discussion of form and function with reference to data from Laos, *Gesture* 1.2, 2001, 185–212 (Chapter 3); Primary and secondary pragmatic functions of pointing gestures, *Journal of Pragmatics* 39, 2007, 1722–1741 (Chapter 4, co-authored with Sotaro Kita and J.P. de Ruiter); On linear segmentation and combinatorics in co-speech gesture: A symmetry-dominance construction in Lao fish trap descriptions, *Semiotica* 149.1/4, 2004, 57–123 (Chapter 5); The body as a cognitive artifact in kinship representations: Hand gesture diagrams by speakers of Lao, *Current Anthropology* 46.1, 2005, 51–81 (Chapter 6); Producing and editing diagrams using co-speech gesture: Spatializing non-spatial relations in explanations of kinship in Laos, *Journal of Linguistic Anthropology* 13.1, 2003, 7–50 (Chapter 7).

For their help in putting the book together, and in commenting on newly written sections, I thank Paul Kockelman, Steve Levinson, and Tanya Stivers. I am grateful to Adam Kendon for encouragement and advice at a critical stage. I owe the ‘pyramid of midgets’ line, with thanks, to Peter Richerson and Rob Boyd (2005: 50).

While I cannot list each of the many people who have contributed to this work, a few deserve special mention. Steve Levinson has supported, encouraged, and contributed to the development of my work for nearly a decade. For this, and for his camaraderie, intellectual and otherwise, I am especially grateful. Sotaro Kita introduced me to gesture research when, in the year 2000, he hand-delivered an entire suitcase full of field equipment to me in Melbourne, thus beginning my career in video-based fieldwork. In subsequent years his constructive approach was a welcome path into the world of gesture. At the same time, and since, J. P. de Ruiter has helped me keep gesture research in its proper perspective. David Wilkins preceded me in Nijmegen, and the indelible traces he left there have been a major influence. My meetings with David have been few and far between, but those who know him will understand how even brief exchanges can have far-reaching effects. Bill Hanks has been a key interlocutor and friend, and a steady advocate for the importance of situated, token understandings, with obvious effect on my work. Also a major influence in recent years is Paul Kockelman, whose riotous intellect has knocked down walls in my understanding of meaning. And last but not least, Tanya Stivers has re-educated the linguist in me, bringing sequence and sociality to the analytic foreground.

This work would not have been possible without the generosity of many Lao speakers who allowed me to videotape them in action and to publicize the results: in particular, I want to thank Mr Naak, Mr Phouthong, Mrs Thong, and Na Vayaphanh.

And finally thanks to family for their patience and support, not only classificatory kin (Edith and the Gang of 5) and affinal kin (Na) at home, but consanguines and their people on the other side of the globe.

N.J.E.

NIJMEGEN

1 Composite utterances

A theory of utterance should not begin with a division between ‘speech’ and ‘gesture’.

Adam Kendon, 1986

In human social behaviour, interactants build communicative sequences, move by move. These moves are never semiotically simple. Their composite nature is widely varied in kind: a word combined with other words, a string of words combined with an intonation contour, a diagram combined with a caption, an icon combined with another icon, a spoken utterance combined with a hand gesture. By what means does an interpreter take multiple signs and draw them together into unified, meaningful packages? This book explores the question with special reference to one of our most familiar types of move, the speech-with-gesture composite.¹

This introductory chapter sketches a view of how it is that interpreters may derive meaning from composite utterances. This view of meaning has emerged from the empirical studies in Chapters 2–7, but is intended to have more general application to the analysis of any kind of communicative move, regardless of whether it involves speech, gesture, both, or neither.

1.1 Meaning does not begin with language

In a person’s vast array of communicative tools, language is surely unrivalled in its expressive richness, speed, productivity, and ease. But the interpretation of linguistic signs is driven by broader principles, principles of rational cognition in social life, principles which underlie other processes of human

¹ Following from a long line of luminaries: De Jorio (2000 [1832]), Wundt (1973 [1921]), Efron (1972 [1941]), Goffman (1963), Condon and Ogston (1967), Ekman and Friesen (1969), Birdwhistell (1970), Kendon (1972, 2004), Slama-Cazacu (1976), Schegloff (1984), McNeill (1985, 2005), Calbris (1990), Haviland (1993, 2000), Streeck (1993, 1994), Goodwin (1994, 2006), Bavelas (1994), Engle (1998), Müller (1998), de Ruiter (2000), Beattie (2003), Goldin-Meadow (2003a), Liddell (2003), Kita and Özyürek (2003), Brookes (2004), Gullberg (2008), among very many others.

judgement, from house-buying to gambling to passing people on a crowded street. So, to understand meaning, we ought not begin with language (Enfield and Levinson 2006: 28). There is meaning in language for the same reason there is meaning elsewhere in our social lives: because we take signs to be public elements of cognitive processes (Peirce 1955), evidence of others' communicative intentions (Grice 1957, 1975). Our clues for figuring out those intentions are found not only in conventional symbols like words, but in the rich iconic–indexical relations which weave threads between just about everything in sight (Peirce 1955, Silverstein 1976, Levinson 1983, Kockelman 2005). Language is just a subset of the full resources necessary for recognizing others' communicative and informative intentions.

1.2 Meaning is dynamic, motivated, and concrete

Among fashions of thinking about language over the last century, a dominant neo-Saussurean view says that meaning is a representational relation of phonological form to conceptual content: a sign has meaning because it specifies a standing-for relation between a signifier and a signified. Semanticians of many stripes agree on this (cf. Jackendoff 1983, Cruse 1986, Langacker 1987, Wierzbicka 1996, among many others). But there is reason to question whether a view of signs as static, arbitrary, and abstract is an adequate depiction of the facts, or even optimal as an analytic framework of convenience. There is reason to stay closer to the source, to see signs as they are, first and foremost: dynamic, motivated, and concrete (Hanks 1990). Standard statements about meaning such as 'the word X means Y' really mean 'people who utter the word X are normatively taken by others to intend Y across a sufficiently broad range of contexts'. We should not, then, understand dichotomies like static versus dynamic, arbitrary versus motivated, or abstract versus concrete as merely two sides of a single coin. The relation is asymmetrical, since we are always anchored in the dynamic–motivated–concrete realm of contextualized communicative signs.

Some traditions doubt whether a Saussurean 'form–meaning mapping' account of meaning is appropriate. In research on co-speech hand gesture, for example, McNeill (2005) has forcefully questioned the adequacy of a coding-for-decoding model of communication. The same point has long been made for more general reasons, in more encompassing theories of semiosis, and in theories of how types of linguistic structure mean what they mean when used as tokens in context (Grice 1975). If we need alternatives to a static view of meaning, general tools are already available for addressing specific problems raised by co-speech gesture. These tools come from two sources: (neo-)Peircean semiotics (e.g. Peirce 1955, Colapietro 1989, Parmentier 1994, Kockelman

2005) and (neo-)Gricean pragmatics (e.g. Grice 1975, Levinson 1983, 2000, Sperber and Wilson 1995 [1986], Horn 1989, Atlas 2005). Subsequent sections explore the relevant analytic tools offered by these traditions.

1.3 Meaning is a composite notion

To set the stage, we anchor the discussion with a few examples of composite signs. Figure 1.1 shows a man kneeling, atop steps, with a crowd looking on.

While the kneeling posture may have an intrinsic, ethological basis for interpretation, this particular token of the behaviour has had a deeply enriched meaning for many who have seen it, because it was performed by this particular man, at this time and place. The man is Willy Brandt, chancellor of



Figure 1.1 Man kneeling atop steps, with crowd looking on.

West Germany. Once you know just this, the act already begins to take on enriched meaning. It is not just a man kneeling, but a man whose actions will be taken to stand for those of a nation's people. It is 7 December 1970, a state visit to Warsaw, Poland. These new layers of information should yet further enrich your interpretation. To add another layer: the occasion is a commemoration of Jewish victims of the Warsaw Ghetto uprising of 1943. Brandt later described the moment: 'On the abyss of German history and carrying the burden of the millions who were murdered, I did what people do when words fail them.' The body posture illustrated in Figure 1.1 is a composite sign in so far as its meaning is partly a function of its co-occurrence with other signs: in particular, the role being played by its producer, given the circumstances of its time and place of production. The behaviour derives its meaning as much from its position on these coordinates as from its intrinsic significance. As Wittgenstein put it, 'Only when one knows the story does one know the significance of the picture' (Wittgenstein 1953: I-§663).

Brandt's *Kniefall* is special partly because it was not accompanied by speech. Most composite utterances, including the speech-with-hand-movement utterances discussed in this book, do include a linguistic component. A relatively simple example of a composite sign with words is the image-with-



Figure 1.2 William Henry Fox Talbot, *Scene in a library*, 1845.

caption format typified by photographs and artwork, as in Figure 1.2. This photograph, titled *Scene in a library*, features wooden shelves with books on them. What makes this a composite sign is that the visual image and the string of words are taken together as part of the artist's single overall intention (Preissler and Bloom 2008, cf. Richert and Lillard 2002). The image and the words are different types of signs, but they are presented together, and taken together, in a composite.

As with any artwork's title, Talbot manipulates our attention to the image. Even if he had given the work a more directly descriptive title like *Books*, this would still invite us to attend differentially to what we actually see. A title *Books* would omit mention of the shelves, in line with the asymmetry in the image (the shelves are not visually foregrounded either). The title Talbot actually used – *Scene in a library* – does not narrow in on any part of the image, in fact it draws our attention to an imagined larger context which is not visible at all. We take the work to represent a scene in a library, and we trustingly presume the photograph to have actually been taken in a library, thanks to the verbal instruction embodied in the work's title. This presumption



Figure 1.3 William-Adolphe Bouguereau, *The wave*, 1896.

is easily made, under a general heuristic of semiotic unity: when encountering multiple signs which are presented together, take them as one.

This presumed bond of word and image is manipulated for wry effect in Figure 1.3.

Strangely resembling a pin-up from an auto magazine ostensibly featuring a new make of car, the title of this image makes a perverse claim about what is being presented. Despite the strong attention-directing force of the nude's blatant centrality in the image (cf. Clark *et al.* 1983), the composite sign's linguistic component directs our attention elsewhere. With the image, Bouguereau gives us a nude, but with the title he purports to depict not a nude, not even a nude by the sea, but a wave.²

These three examples illustrate essentially the same phenomenon as we find in the co-occurrence of expressive hand movements with speech: context-situated composites of multiple signs, part conventional, part non-conventional. Compare them to Figure 1.4, an image from a video-recording showing three Lao men sitting in a village temple, one of them thrusting his arm forward and down, with his gaze fixed on it.

The discussion is about construction works underway in the temple. The man on the left is reporting on a problem in the installation of drainage pipes from a bathroom block. He says that the drainage pipes have been fixed at too shallow an angle, and they should, instead, drop more sharply, to ensure good run-off. As he says *haj5 man2 san2 cang1 sii4* 'Make it steep like this', he thrusts his arm forward and down, fixing his gaze on it, as shown in Figure 1.4. The meanings of his words and his gesture are tightly linked, through at least three devices: (1) their tight spatiotemporal co-occurrence in place and time (both produced by the same source), (2) the use of the explicit deictic expression 'like this' (sending listeners on a search: 'Like what?'), (3) the use of eye gaze for directing attention.

A similar case is presented in Figure 1.5, from a description of a type of traditional Lao fish trap called the *sòdn5* (see Chapter 5).

Again we see a speaker's overall utterance meaning as a unified product of multiple sources of information: (a) a string of words (itself a composite sign consisting of words and grammatical constructions), (b) a two-handed gesture, (c) tight spatiotemporal co-occurrence of the words and gestures (from a single source), and (d) eye gaze directed toward the hands, also helping to connect the composite utterance's multiple parts. This is subtly different from Figure 1.4 in that it does not involve an explicit deictic element in the speech (cf. the overt 'like this' element in Figure 1.4 which obliges us to consult the

² He may of course also be inviting us to find qualities in common between a wave and the human figure depicted.



Figure 1.4 Man (left of image) speaking of preferred angle of a drainage pipe under construction: ‘Make it steep like this.’

gesture to complete the utterance’s meaning). Like the image-with-caption examples in Figures 1.1–1.3, spatiotemporal co-placement in Figure 1.5 is sufficient to signal semiotic unity. The gesture, gaze, and speech components of the utterance are taken together as a unified whole. As interpreters, we effortlessly integrate them as relating to one overall idea.³

A general theory of composite meaning takes Figures 1.1–1.5 to be instances of a single phenomenon: signs co-occurring with other signs, acquiring unified meaning through being interpreted as co-relevant parts of a single whole. A general account for how the meanings of multiple signs are unified in any one of these cases should apply to them all, along with many other species of composite sign, including co-occurring icons in street signs, grammatical unification of lexical items and constructions, and speech-with-gesture composites.

In studying speech-with-gesture, we should register two desiderata for an account of composite meaning. A first requirement is to provide a modality-

³ As Freud argued, with many since, there may be leakage of unintended, apparently unrelated information, particularly through modalities over which a sender has less control.



Figure 1.5 Man describing the *sòòn5*, a traditional Lao fish trap: ‘As for the *sòòn5*, they make it fluted at the mouth.’ (See Chapter 5.)

independent account of ‘gesture’ (Okrent 2002). While we want to capture the intuition that co-speech hand gesture (manual–visual) conveys meaning somehow differently to speech (vocal–aural), this has to be articulated without reference to modality. We need to be able to say what makes speech-accompanying hand movements ‘gestural’ in such a way that we can sensibly ask as to the functional equivalent of co-speech gesture in other kinds of composite utterances; for example, in sign language of the Deaf (all visual, but not all ‘gesture’), or in speech heard over the phone (all vocal–aural, but not all ‘language’).

A second desideratum for an account of meaning in speech-with-gesture composites is to capture the notion of ‘holistic’ meaning in hand gestures, the idea that a hand gesture has the meaning it has only because of the role it plays in the meaning of an utterance as a whole (McNeill 1992, 2005, Engle 1998). Consistent with an aim for analytic generality, I argue that a notion of holistic meaning is required not only for analysing the meaning of co-speech hand gesture, but more generally for analysing linguistic and other types of signs as well (including wordless moves like Brandt’s *Kniefall*). This results from acknowledging that an interpreter’s task begins with the recognition of a signer’s communicative intention (i.e. recognizing that the signer has an informative intention). The subsequent quest to lock onto a target informative

intention can drive the understanding of the composite utterance's parts, and not necessarily the other way around.

1.4 The anatomy of meaning in composite utterances

1.4.1 Contexts of hand gesture

One view of speech-with-gesture composites is that the relation between co-expressive hand and word is a reciprocal one: 'the gestural component and the spoken component **interact** with one another to create a precise and vivid understanding' (Kendon 2004: 174, original emphasis; cf. Özyürek *et al.* 2007). By what mechanism does this reciprocal interaction between hand and word unfold? Different approaches to analysing meanings of co-speech gestures find evidence of a gesture's meaning in a range of sources, including (i) speech (coterminous) which co-occurs with the hand movement, (ii) a (prior) stimulus or cause of the utterance in which the gesture occurs, (iii) a (subsequent) response to, or effect of, the utterance, or (iv) purely formal characteristics of the gesture. These four sources (often combined) draw on different components of a single underlying model of the communicative move and its sequential context, where the hand-movement component of the composite utterance is contextualized from three angles: A. what just happened; B. what else is happening now; C. what happens next. This is illustrated in Figure 1.6.

The three-part sequential structure illustrated in Figure 1.6 underlies a basic trajectory model recognized by many students of human social behaviour. Schutz (1970), for example, speaks of actions (at B) having 'because motives' (at A) and 'in-order-to motives' (at C; e.g. 'I'm picking berries [B] because I'm hungry [A], in order to eat them [C]'; cf. Sacks 1992, Schegloff 2007b among many others).

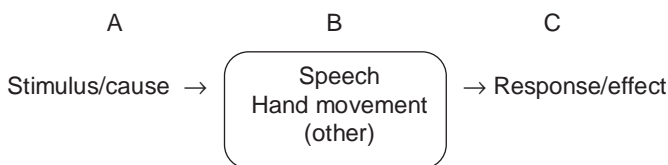


Figure 1.6 Three contexts of hand movement, in sequential interaction: at B, composite utterances may include multiple simultaneous signs; a preceding stimulus/cause at A determines a sign's **appropriateness**; a response determines its **effectiveness**.

1.4.2 *Enchrony: an underlying ontology for the context of composite utterances*

The structure in Figure 1.6 directs our attention to an ontology of the composite utterance as a situated unit of social behaviour with causes (or conditions) and effects (Goffman 1964, Schegloff 1968). An intentional cause and interpretive effect are as definitive of the process of meaning as the pivotal signifying behaviour itself. Any communicative move may be seen as arising more or less appropriately from certain commitments and entitlements, and in turn bringing about new commitments and entitlements (Austin 1962, Searle 1969), for which interlocutors are subsequently accountable. As an analytical framework, this remedies the static, decontextualized nature of Saussure's version of meaning (Kockelman 2005). But this is not merely because it recognizes that meaning arises through a process (McNeill 2005), it is because it recognizes the causal/conditional and normative anatomy of sequences of communicative interaction, where each step brings about a new horizon, with consequences for the people involved (Schegloff 1968, Sacks *et al.* 1974, Goffman 1981, Heritage and Atkinson 1984). Accordingly, we need a term for a causal, dynamic perspective on language whose granularity matches the pace of our most experience-near, moment-by-moment deployment of utterances, not historical time (for which the term *diachronic* is standard) but conversational time. For this I invent the word *enchronic*.⁴ While diachronic analysis is concerned with relations between data from different years (with no specified type or directness of causal/conditional relations), enchronic analysis is concerned with relations between data from neighbouring moments, adjacent units of behaviour in locally coherent communicative sequences (typically, conversations). McNeill (2005) uses *epigenesis* for the real-time birth and development of a composite utterance from a producer's point of view. This is distinct from the intended meaning of enchronic here, namely the intersection of (a) a social causal/conditionality of related signs in sequences of social interaction and (b) a particular level of temporal granularity in a conditionally sequential view of language: conversational time. An enchronic perspective adopts the sequential analytic approach whose application in empirical work was pioneered by Schegloff (1968) and Sacks (1992), following earlier work in sociology. To call it enchronic rather than merely sequential (in the technical sense of Schegloff 2007b) draws attention to the broader set of alternative viewpoints on

⁴ This is an adjective, whose nominal form is *enchrony*. The prefix *en-* refers elsewhere to causal/conditional relations and to the notions of increment and change of state (e.g. *endear*, *enfold*, *enliven*, *enrich*, *encage*).

systems and processes of meaning which we often need to switch between (phylogenetic, diachronic, ontogenetic, synchronic).

1.4.3 *The move: a basic-level unit for social interaction*

An enchronic perspective takes as a primitive unit the communicative move (Goffman 1981). A move may be defined as a recognizable unit contribution of communicative behaviour constituting a single, complete pushing forward of an interactional sequence by means of making some relevant social action recognizable (e.g. requesting the salt, passing it, saying *Thanks*). In communication, a richly multimodal flux of impressions is brought to order by these joint-attentional pulses of addressed behaviour (e.g. bursts of talk) marked off in the flow of time and space, yielding sequences of co-contingent social action (Goodwin 2000a, Schegloff 2007b). The linguistic utterance is a well-studied (if idealized)⁵ type of instantiation of the move (cf. Austin 1962, Searle 1969). With this basic-level status, the linguistic move will be homologous with usage-based analytic units of language such as the clause (Foley and Van Valin 1984), the intonation unit (Pawley and Syder 2000, Chafe 1994), the turn-constructive unit (Sacks *et al.* 1974), the growth point (McNeill 1992), the composite signal (Engle 1998, cf. Clark 1996), and the utterance as multimodal ensemble (Kendon 2004, Goodwin 2000a). But whatever its physical form, the move is a single-serve vehicle for effecting action socially.

An important argument in favour of the move's primitive or basic-level status is its role in the acquisition of communicative skills in children. Before learning their first words, children master the move, beginning with its prototype, the pointing gesture (Kita 2003). A line of research in developmental psychology has identified the onset of the pointing gesture as a watershed moment in the development of human social cognitive and communicative capacities, both ontogenetically and phylogenetically (Bates *et al.* 1975, 1987, Liszkowski *et al.* 2004, Tomasello 2006). The pointing gesture is mastered by prelinguistic infants (at around twelve months of age) and is the first type of move to unequivocally display the sort of shared intentionality unique to human communication and social cognition (Tomasello *et al.* 2005, Liszkowski 2006, Frith and Frith 2007).

The move is therefore a starting point, a seed, a template for the deployment of signs in interaction. On the one hand, the move is a brick for larger structures, building up and out, into conversational sequences and other kinds of coherent discourse structure (Halliday and Hasan 1976, Schegloff 2007b).

⁵ I say it is idealized because there is always more than just language available for an interpreter.

On the other hand, it is a frame or exoskeleton within which internal semiotic complexity may appear, building down and in, yielding phrase distinctions, morphosyntax, information structure, and logical semantics. The work of this book is to examine the kinds of structure that arise when moves are built from word and hand together, and we shall witness both their internal elaboration (down-and-in) and their role in structuring higher-order sequences (up-and-out).

1.4.4 *Conventional and non-conventional components of composite utterances*

Three types of sign are important in interpreting composite utterances: conventional signs, non-conventional signs, and symbolic indexicals.⁶ A conventional sign is found when people take a certain signifier to stand for a certain signified because that is what members of their community normatively do (Saussure 1959 [1916]).⁷ This kind of sign allows for arbitrary relations like /k^hæt/ referring to ‘cat’, by which the cause of my taking [k^hæt] to mean ‘cat’ is my experience with previous occasions of use of tokens of the signifier /k^hæt/. Examples of conventional signs include words and grammatical constructions, idioms, and ‘emblem’ hand gestures such as the OK sign, V for victory, or the finger (Ekman and Friesen 1969, Brookes 2004). Non-conventional signs, by contrast, are found when people take certain signifiers to stand for certain signifieds not because of previous experience with that particular form–meaning pair or from social convention, but where the standing-for relation between form and meaning comes about by virtue of just that singular event of interpretation. Examples include representational hand gestures (in the sense of Kita 2000), that is, where the gesture component of an utterance is a token, analogue representation of its object.

The symbolic indexical is a hybrid of the two types of sign just described, having properties of both.⁸ These include anything that comes under the rubric of deixis (Fillmore 1997 [1971], Levinson 1983), that is, form–

⁶ For convenience, I simplify the analysis of sign types employed here. A full anatomy of sign types would lay out the logical possibilities first mapped by Peirce (1955), and most accessibly interpreted by Parmentier (1994) and Kockelman (2005). The notion of conventional sign here corresponds to Peirce’s symbol; non-conventional sign includes his icon and index. The Peircean type/token distinction (Hutton 1990) cuts across these (see below).

⁷ By saying that a behaviour is normative, I mean that carrying out the behaviour under the appropriate conditions will be effective, and will not evince justified surprise or sanction (cf. Brandom 1979, Kockelman 2006).

⁸ Because symbolic indexicals have both conventional and non-conventional components, statements about those kinds of signs also hold for the relevant components of symbolic indexicals.

meaning mappings whose proper interpretation depends partly on convention and partly on context (Bühler 1982 [1934], Jakobson 1971, Silverstein 1976). Take for example *him* in *Take a photo of him*. Your understanding of *him* will depend partly on your recognition of a conventional, context-independent meaning of the English form *him* (third person, singular, male, accusative) and partly on non-conventional facts unique to the speech event (e.g. whichever male referent is most salient given our current joint attention or common ground). Symbolic indexicals play a critical role in many types of composite utterance, since their job is to glue things together, including words, gestures, and (imagined) things in the world (see Part I of this book).

In the context of these three kinds of sign, it is important to be mindful of the distinction between type and token (Peirce 1955, Hutton 1990). All of the signs discussed above occur as tokens, that is, as perceptible, contextualized, unique instances. But only conventional signs (including conventional components of symbolic indexicals) necessarily have both type and token identities. That is, when they occur as tokens, they are **tokens of types**, or what Peirce called replicas. It is because of their abstract type identity that conventional signs can be regarded as meaningful independent of context, as having ‘sense’ (Frege 1960 [1892]), ‘timeless meaning’ (Grice 1989) or ‘semantic invariance’ (Wierzbicka 1985, 1996). Conventional signs are pre-fabricated signs, already signs by their very nature. They have ‘I am an addressed, relevant sign’ stamped on their foreheads. By contrast, non-conventional signs (including non-conventional components of symbolic indexicals) are tokens but not tokens of types. They are singularities (Kockelman 2005). They become signs only **when taken as signs in context**. This is the key to understanding the asymmetries we observe in composite utterances like speech-with-gesture ensembles. A hand gesture may be a conventional sign (e.g. as ‘emblem’). Or it may be non-conventional, only becoming a sign because of how it is used in that context (e.g. as ‘iconic’ or ‘metaphoric’). Or it may be a symbolic indexical (e.g. as pointing gesture, with conventionally recognizable form, but dependent on token context for referential resolution). Hand gestures are not at all unique in this regard: the linguistic component of an utterance may, similarly, be conventional (e.g. words, grammar), non-conventional (e.g. voice quality, sound stretches), or symbolic indexical (e.g. demonstratives like *yay* or *this*). Ditto for sign components of graphs, diagrams, and other illustrations. Sensory or articulatory modality is no obstacle to semiotic flexibility.

Before concluding this chapter, it is worthwhile registering a common inconsistency in discussion of the meaning of hand movements in composite utterances. The problem is an asymmetric treatment of the way meaning is attributed to words, on the one hand, and gestures, on the other. Linguistic

items like words are often described merely in terms of what they conventionally **encode** (as standing for lexical **types**), while gestures are typically described in terms of what they non-conventionally **convey** (as standing for utterance-level **tokens** of informative intention). In other words, the interpreter's problem of comprehending word meaning is taken to be one of **recognition** (from token form to type lexical entry), while the problem of comprehending gesture meaning is taken to be one of **interpretation** (from token form to token informative intention). The inconsistency here is that it overlooks the fact that comprehension of the linguistic component **also** involves interpretation, yielding token informative intentions. In interpreting the meanings of words, we do not stop with mere recognition of type lexical entries, but, just like with gestures, we **also** use them for recognizing a speaker's token informative intention. To illustrate, take an example cited by McNeill (2005: 26), in which a speaker says *and he came out the pipe* while doing an 'up-and-down away' hand gesture (the hand is moving away from the body as it is moved repeatedly up and down). Hearing *came out*, an interpreter recognizes these sounds to be tokens of types (i.e. with the meaning 'came out'). He or she may also enrich this meaning 'came out' in using it as a clue for figuring out the speaker's informative intention in producing this composite utterance. They may of course exploit the accompanying gesture in this process of enrichment. In the experiment described by McNeill, a subject who heard the first speaker's description of the scene as *and he came out the pipe*_[GESTUREup-and-down-away] later re-describes it as *the cat bounces out the pipe*.⁹ This shows that **both** the gesture and the words are enriched by their co-occurrence in that context, being taken to be co-occurring signs of a single informative intention. *Came out* and _[GESTUREup-and-down-away] together point to a single idea 'bounces out'. While word recognition has no analogue in the interpretation of the iconic gesture (since the gesture is a token but not a token of a type), attribution of overall utterance-intention of words **does** have an analogue in the interpretation of the gesture. When examining gesture, as when examining any other component of composite utterances, we must carefully distinguish between token meaning (enriched, context-situated), type meaning (raw, context-independent, pre-packaged), and sheer form (no necessary meaning at all outside of a particular context in which it is taken to have meaning). These distinctions may apply to signs in any modality.

⁹ Note that the re-teller not only enriches *came out*_[GESTURE up-and-down-away] as 'bounces out', he also enriches *he* as 'the cat'; regarding the pronoun *he* in the original utterance, the subject must have both recognized *he* as a token of the type 'he', which stands in this case for a token informative intention 'the cat'.

-
-
- I. Encoded
 - I.1. Lexical (open class, symbolic)
 - I.2. Grammatical (closed class, symbolic indexical)
 - II. Enriched
 - II.1. Indexical resolution
 - II.1.1 Explicit (via symbolic indexicals, e.g., pointing or demonstratives)
 - II.1.2 Implicit (e.g., from physical situation)
 - II.2. Implicature
 - II.2.1 From code
 - II.2.2 From context
-
-

Figure 1.7 Sources of composite meaning for interpretation of communicative moves. ‘Encoded’ = conventional sign components. ‘Enriched’ = non-conventional token meanings drawing on context.^{10, 11}

1.4.5 *Elements of composite utterances*

Based on the discussion so far, we may define the composite utterance as a communicative move that incorporates multiple signs of multiple types. Sources of these types of sign are given in Figure 1.7 (cf. Levinson 1983: 14, 131, Hanks 1990: 51ff.).

Composite utterances are interpreted through the recognition and bringing together of these multiple signs under a pragmatic unity heuristic or co-relevance principle, i.e. an interpreter’s steadfast presumption of pragmatic unity despite semiotic complexity.

¹⁰ Encoded meaning encompasses both lexical and grammatical meaning. Grammatical signs show greater indexicality because they signify context-specific ties between two or more elements of a composite utterance (e.g. grammatical agreement, case-marking etc.) or between the speech event and a narrated event (Jakobson 1971; e.g. through tense-marking, spatial deixis etc.).

¹¹ Indexical enrichment refers to the resolution of reference left open either explicitly (e.g. through symbolic indexicals like *this*) or implicitly (e.g. by simple co-placement in space or time; thus, a ‘no smoking’ sign need not specify ‘no smoking **here**’). Enrichment through implicature refers to Gricean token understandings, arising either through rational interpretation based on knowledge of a restricted system of code (i.e. informativeness scales and other mechanisms for Generalized Conversational Implicature; cf. Levinson 2000), or through rational interpretation based on cultural or personal common ground (e.g. Particularized Conversational Implicatures such as those based on a maxim of relevance; Sperber and Wilson 1995 [1986]).

1.5 Sign filtration: triggers and heuristics

The taxonomy of elements of composite signs in Figure 1.7 presupposes that an interpreter can solve the problem of sign filtration, i.e. that they can parse out from a flux of impressions those things that are to be taken as signs. This filtration is assisted by triggers which direct us to lock on to certain signs, constraining the search space. An important trigger is that a perceptible impression must be recognizable as addressed, that is, being produced by a person for the sake of its interpretation by another. Conventional signs like words have this addressed-ness by their very nature. But other perceptibles are only potential signs, and their addressed-ness needs to be specially marked. This can be achieved by means of attention-drawing indexicals (hand pointing, saying ‘like this’ etc.), by sheer spatiotemporal co-occurrence, or by special diacritic marking (see Figures 1.1–1.5, above). An example of the latter is discussed in Chapter 3, where movements of the face and head can serve as triggers for eye gaze to be interpreted as pointing, not merely as looking. In yet other cases, interpreters can employ abductive, rational interpretation to detect that an action is done with a communicative intention (Peirce 1955, Grice 1957). For instance, if you open a jar I may be unlikely to take this to be communicative, but if you carry out the same physical action without a jar in your hands, the lack of conceivable practical aim is likely to act as a trigger for implicature (Levinson 1983: 157, Gergely *et al.* 2002).

The data in this book do not present particular difficulties for interpreters in detecting communicative intention or identifying which signs to include when interpreting a composite utterance. Mostly, the mere fact of language being used triggers a process of interpretation, and the gestures which accompany speech are straightforwardly taken to be associated with what a speaker is saying (Kendon 2004). Hand gestures are therefore available for inclusion in a unified interpretation, whether or not we take them to be intended to communicate.

Note the kinds of heuristics that are likely being used in solving the problem of sign filtration. By a **convention heuristic**, if a form is recognizable as a socially conventionalized type of sign, assume it stands for its socially conventional meaning. Symbols like words may thus be considered as pre-fabricated semiotic processes: their very existence is due to their role in communication (unlike iconic–indexical relations which may exist in the absence of interpretants). By an **orientation heuristic**, if a signer is bodily oriented toward you, most obviously by body position and eye gaze, assume they are addressing you. By a **contextual association heuristic**, if two signs are contextually associated, assume they are part of one signifying action. Triggers for contextual association are timing and other types of indexical proximity (e.g. placing caption and picture together, placing word and gesture

together). By a **unified utterance–meaning heuristic**, assume that contextually associated signs point to a unified, single, addressed utterance–meaning. And by an **agency heuristic**, if a signer has greater control over a behaviour, assume (all things being equal) that this sign is more likely to have been communicatively intended. Language scores higher than gesture on a range of measures of agency (Kockelman 2007).

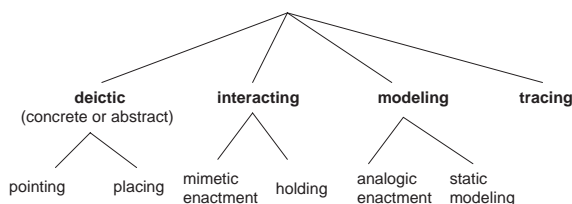
1.6 Semiotic analysis of gestures

Like any signs, hand movements can stand for things in three essential ways (often in combination), referred to by Peirce (1955) as types of ground: iconic, indexical, symbolic. These crucial yet widely mishandled distinctions are defined as follows. A relation of a sign standing for an object is iconic when the sign is taken to stand for the object because it has perceptible qualities in common with it. The sign is indexical when it is taken to stand for an object because it has a relation of actual contiguity (spatial, temporal, or causal) with that object. The relation is symbolic when the sign is taken to stand for an object because of a norm in the community that this sign shall be taken to stand for this object. These three types of ground are not exclusive, but co-occur. A fingerprint on the murder weapon is iconic and indexical. It is iconic in that it has qualities in common with the pattern on the killer's actual fingertip and in this way it is a sign that can be taken to stand for the fingertip. It is indexical in that (a) it was directly caused by the fingertip making an impression on the weapon (thus a sign standing for an event of handling it), and (b) the fingertip of the killer is in contiguity with the whole killer (thus a sign standing for the killer himself). Standard taxonomies of gesture types (McNeill 1992, Kendon 2004, *inter alia*) are fully explicable in terms of these types of semiotic ground (Figure 1.8).¹²

An exhaustive analysis of the semiotics of hand gestures will need to systematically explore their values on the many parameters along which signs differ: Formal segmentability, stability across populations, evanescence or persistence in time from production, symmetry of perceptual access for producer and interpreter, relative immediacy of the processes of production and interpretation, portability, combinatorics, information structure (cf. Kockelman 2005: 240–241). This will entail teasing apart the large set of distinct semiotic dimensions which hand movements incorporate (Talmy 2006).¹³ Hand

¹² Figure 1.8 presents only some types of hand gesture of interest in this book. See references in note 1, above, for discussion of these and numerous other types such as 'beats', 'adaptors', and 'quotable gestures'.

¹³ The notion of semiotic dimension is distinct from that of sensory modality. Any sign medium within a single sensory modality may have numerous distinct semiotic dimensions, meaning any distinguishable, simultaneously variable perceptible component of a sign vehicle which

**Deictic:**

- semiotic function: *indexical* (in that the directional orientation of the gesture is determined by the conceived location of a referent), and *symbolic* (in that the form of pointing can be locally conventionalized); the hands are used to bring the referent and the attention of the addressee together;
 - in *concrete deixis*, the referent is a physical entity in the speech situation, while in *abstract deixis* the referent is a reference-assigned chunk of space with stable coordinates
 - in *pointing*, the attention of the addressee is directed to the referent by some vector-projecting articulator (such as the index finger or gaze)
 - in *placing*, the referent is positioned for the attention of the addressee
(Nb.: *Gaze* plays an important role in deictic gestures; it projects its own attention-directing vector which may (a) reinforce a deictic hand gesture by providing a second vector oriented towards the same referent, and (b) assist in the management of attention direction during production of other gestures.)

Interacting:

- semiotic function: *iconic* (in that the hands imitate an action) and *indexical* (in that the shape of the hands is not the shape of the referent, but is *determined by* the shape of the referent); the hands are meant to look as if they were interacting with the referent;
 - in *mimetic enactment*, the hands are moving as if they are doing something to or with the referent
 - in *holding*, the hands are shaped to look as if they are holding the referent

Modeling:

- semiotic function: *iconic*; the hands are meant to look as if they are the referent
 - in *analogic enactment*, the hand's movement imitates the movement of the referent
 - in *static modeling*, the hand's shape imitates the shape of the referent

Tracing:

- semiotic function: *iconic* (in that the gesture imitates drawing) and *indexical* (in that only part of the referent is depicted, but the whole is referred to); the hands (more specifically, the fingers) are meant to look as if they were tracing the shape of some salient feature of the referent, such as its outline.
-
-

Figure 1.8 Some semiotic devices used in illustrative co-speech gestures discussed in this book (cf. Mandel 1977, Kendon 1988, Müller 1998).

movements are well suited to iconic–indexical meaning thanks to their rich potential for sharing perceptible qualities in common with physical objects and events. But they are not at all confined to these types of meaning. As Wilkins writes, '[the] analog and suprasegmental or synthetic nature [of gestures] does not make them any less subject to convention, and does not deny them combinatorial constraints or rules of structural form' (Wilkins 2006: 132). For example, in some communities, 'the demonstration of the

could conceivably be taken to be a sign for something (de Ruiter *et al.* 2003). For example, upon uttering a word, the human voice can simultaneously vary many distinct features of a speaker's identity (sex, age, origin, state of arousal, individual identity etc.), along with pitch, loudness, among other things. What makes pitch and loudness distinct semiotic dimensions is that they can be varied independently of each other. But loudness is a single dimension, because it is impossible to produce a word simultaneously at two different volumes.



Figure 1.9 Man shows the size of a fish using his forearm as a measure. Note his eye gaze as a cue for the communicative relevance of what his hand is doing (cf. Figures 1.4, 1.5, above)

length of something with two outstretched hands may require a flat hand for the length of objects with volume (like a beam of wood) and the extended index fingers for the length of essentially linear objects lacking significant volume (e.g. string or wire)’ (ibid.). A similar example is the Lao speaker’s conventional way of talking about sizes of fish, illustrated in Figure 1.9, by using the hand or hands to encircle a cross-section of a tapering tubular body part such as the forearm, calf, or thigh. This is taken as standing for the actual size of a cross-section of the fish.

Another kind of conventionality in gestures concerns types of communicative practice like, say, ‘tracing’ in mid air (Mandel 1977, Kendon 1988). It may be argued that there are conventions which allow interpreters to recognize that a person is doing an illustrative tracing gesture, based presumably on formal distinctions in types of hand movement in combination with attention-directing eye gaze toward the gesture space (see Part II of this book). While the exact form of a tracing gesture cannot be pre-specified, its general manner of execution may be sufficient to signal **that** it is a tracing gesture.

1.7 Overview of the book

While each chapter deals with composite utterances, Part I focuses on deictic or symbolic indexical components, while Part II focuses on non-conventional illustrative components. Part I concentrates on signs whose main job it is to

link conventional signs with non-conventional signs, while Part II concentrates on signs whose main job it is to create new signs as (virtual) things in the world which may be talked about and pointed to.

The focus of Part I is the role in composite utterances of symbolic indexicals in attaching conventional signs (in most cases, words) to things in the world. In Chapter 2, the symbolic indexicals in question are in the spoken component of the composite utterance, in the form of demonstratives (words like *this* and *that*). In Lao, as in all languages, demonstratives form a closed set, where the members of this set play off each other in how they direct attention to things in context, and express stances toward those things. Conventional signs in a closed set of this kind not only have intrinsic, coded meanings, but also pick up enriched meanings through markedness relations within a restricted paradigm. The Lao system of demonstratives is not different in this respect to the system of pointing discussed in Chapters 3 and 4. The systematic contrast among types of bodily pointing (lip-pointing, ‘small’ finger-points, ‘big’ finger-points) demonstrate that the idea of a system of meaningful oppositions is as applicable to types of hand gestures as it is to types of words. In this sense, gestures can have grammatical properties.

Part II examines composite utterances in which hand gestures are employed to create virtual illustrations which function like models of concrete objects and diagrams of abstract sets of relations. The analyses show speakers using hand movements iconically and indexically to create novel signs which show remarkable structural persistence, both in space and through time. As these structures are established in the common field of attention, pointing plays a crucial role. The structures that the modelling and diagramming gestures create are effectively treated like physical objects. As Chapter 7 shows, speakers have to cope with the consequences of this virtual reality, being required to make explicit editing manoeuvres upon otherwise ethereal structures.

There are some principled differences between the phenomena described in Parts I and II. For example, while the deictic signs in Part I constitute closed grammatical systems, the illustrative signs in Part II are novel, open class items. And while Part I focuses more on the internal structure of composite utterance units (regarding the move as exoskeleton, fleshed out down-and-in), Part II explores complex structures which emerge through sequences of multiple moves (regarding the move as brick, building up-and-out).

These differences in focus of the two parts are offset by some thorough-going common themes. Most important is the collaborative, public, socially strategic nature of the process of constructing composite utterances. These communicative moves are not merely designed but designed for, and with, anticipated interpreters. They are not merely indices of cognitive processes,

they constitute cognitive processes. They are distributed, publicized, and intersubjectively grounded. Each type of composite utterance discussed in this book is regulated by its producer's aim not just to convey some meaning but to bring about a desired understanding in a social other. So, like all instruments of meaning, these composites are not bipolar form–meaning mappings, or mere word-to-world glue, they are premised on a triadic, cooperative activity consisting of you, me, and what I'm trying to say.

P A R T I

Deictic components of moves

The arrow points only in the application that a living being makes of it.

Ludwig Wittgenstein, 1953

2 Demonstratives

‘Here-now’ is never a sheer physical reality.

William F. Hanks, 1990

Words like *this* and *that* play a key role in composite utterances, like glue for sticking the linguistic system onto the physical world. But their semantics remain poorly understood, despite advances in typological description (Fillmore 1982: 47ff., Anderson and Keenan 1985, Himmelmann 1996, Diessel 1999, Dixon 2003), primarily because demonstratives have seldom been examined in situ. Due to the inherently context-bound character of demonstratives, it is necessary to examine their use in spontaneous interaction, in all its richness (Hanks 1990). This is the method pursued in this chapter, using data from video recordings of natural interaction.

The aim of this chapter is to present a case study of the kind of resources a language can offer in guiding an interpreter’s unification of word and world in utterance comprehension. The Lao demonstrative determiner system is a typologically unremarkable two-term system (*nii4* versus *nan4*) opposing so-called proximal and distal (Hasan 1968, Halliday and Hasan 1976, Anderson and Keenan 1985, *inter alia*). Despite this apparent simplicity, a traditional analysis – assuming a symmetrical opposition of ‘distance’ marking between the two terms – fails. Close attention to distributional facts and the pragmatics of interlocutors’ interpretations of physical space in interaction (including contingent factors like attention, common ground, cultural and personal conceptions of space) supports a lean semantic analysis of the two Lao demonstratives, whereby neither literally encodes distance (i.e. neither makes specification of notions such as ‘near’ or ‘far’), and only one encodes location (namely the semantically more specific ‘distal’ demonstrative, which refers to something ‘not here’). The proposed semantics are minimal, yet they remain consistent with the use of these forms in rich contexts. Similar analyses are likely to hold for other such typologically unremarkable systems (such as English *this* and *that*).

While it may be that only a minimal proportion of a given utterance’s whole meaning is semantically encoded, one cannot verify the content of that encoded component nor understand the principles of its use without studying

contingent details of interactional context, monitoring input of information from the speech situation, and resultant enrichment of encoded meaning by inference. In order to understand what is conveyed in total in an interaction, a Gricean distinction between encoded semantics and defeasible implicature is necessary. Maintaining such a distinction helps to highlight the importance of context-given information and context-derived inference in the overall meanings of linguistic utterances.

2.1 Meaning, context, and the semantics/pragmatics distinction

The analysis to be presented here demands the recognition of a distinction between effectively context-independent encoded meanings for linguistic signs, on the one hand, and qualitatively distinct contextual information and context-specific inferences, on the other. This view puts context, culture, and cognition at the heart of linguistic meaning, yet its merits are widely overlooked. It is thus necessary to preface the analysis with a few remarks on theoretical background.

Meaning arises when a perceptible artefact, such as a sight or a sound, is interpreted by an observer to stand for something (Peirce 1965 [1932]: 135). The relation of standing for is not inherent in a signal, but is actively brought about by an interpreting individual. From exposure to some sign *p*, an interpreter may take this to stand for *q*, either because *q* has qualities in common with *p* (an iconic relation), because *q* is in spatial–temporal–causal contiguity with *p* (an indexical relation), or because of a predetermined social agreement that *q* will be taken to stand for *p* (a symbolic relation; Peirce 1965 [1932]: 143). These distinctions do not denote types of signs, but manners of signification. Individual signs often combine more than one.

Let us consider symbolic meaning. A sign is a symbol if its relation of standing for is determined by a precedent of social agreement, and not by any relationship of formal likeness or ‘dynamical connection’ between form and meaning.¹ I use the terms intension and semantic meaning both to refer to the symbolic meaning that a linguistic sign conventionally encodes. The semantic invariant of a linguistic sign is the minimal meaning that is always derived from the signal regardless of the context in which it is used. Some researchers of meaning have expressed unease with this view, claiming it implies that linguistic signs are ‘objectively real entities’ (Lakoff 1987: 171). But an

¹ This does not mean that the form of a symbol cannot have some quality in common with its meaning. Consider onomatopoeic expressions like *cock-a-doodle-doo*. In this case, the form–meaning relationship holds not **because of** a formal likeness between the phonological string and the conceived sound of a rooster crowing, but because of a social agreement that just this string will have just this meaning. This is why the sound–meaning pairing is not predictable (despite being ‘motivated’ – cf. French *cocorico*).

analytical commitment to the idea of encoded symbolic meaning does not entail a view that a sign has a Platonic reality independent of those individuals who hold the idea of that sign. Peirce observed that ‘the word lives in the minds of those who use it’ (Peirce 1965 [1932]: 169) and repeatedly asserted that meaning cannot arise without an active mind bringing it about. Nevertheless, it is **effectively** the case that linguistic signs have stable and context-independent meanings. This is explained by the mechanism of convention, which may be defined as ‘a community’s solution to a recurrent coordination problem’ (Clark 1996: 70; after Schelling 1960, Lewis 1969). The coordination problem relevant to the present discussion is the need for interlocutors – who lack direct access to each other’s minds – to reliably converge on mutual understandings of each other’s intentions and interpretations in talk (Clark 1996). An individual’s representation of a linguistic meaning is a private hypothesis, but the public contract of convention results in the effective convergence, across a population, of individuals’ personal hypotheses about the concept encoded in a word.² This ‘intercalibration of idiolects’ (Hockett 1987: 106–107, 157–158; cf. Lee 1996: 227–228) allows us to trust that meanings will be stable across contexts, to predict their effects, and to be confident of what will and will not go on record when a given sign is used. The public stability of a semantic representation is not independent of human interaction, but arises from a large-scale aggregation of community-wide ‘joint action’ (as described by Clark 1996: 59ff.).

The domain of pragmatics concerns the process whereby meanings richer than the encoded conventional meanings of signs arise in real contexts (Levinson 1983). Pragmatic meanings are semiotically accessible and/or logically derivable, but are not semantically encoded. They include contingent aspects of context and common ground (things one knows, and things one can see and hear), on the one hand, and derived inferences on the other. Context and common ground are input for deriving inferences from encoded meanings. Two people’s common ground is defined as ‘the sum of their mutual, common, or joint knowledge, belief, and suppositions’ (Clark 1996: 93). This knowledge, shared and mutually known to be shared among interlocutors, may be freely assumed for the purpose of calculating inferences. The term implicature refers to the process or product of such inferences. Implicatures are contingent and context-dependent, arising from given common ground, including speakers’ knowledge of the linguistic system, and associated expectations of what a speaker could have said but didn’t (Grice 1975, Levinson 2000). While a linguistic sign may have one implicature in one

² Language acquisition involves hypothesis construction (cf. Brown 1958, Clark 1999, Tomasello 2003), and in adult life our established models of linguistic meanings remain hypotheses, which are repeatedly tested (and potentially revised) whenever we utter or hear linguistic signs.

context, this implicature may simply not arise in another context. Semantic meaning, by contrast, is derived regardless of context. A final relevant notion under the rubric of pragmatics is extension (in the sense of Carnap 1947). The extension of a linguistic sign is the array of things, events, or situations to which it can be taken to refer.

The insight that meaning is derived from linguistic signs in a combination of qualitatively distinct ways has led to the most important working distinction for the analysis of how linguistic forms are used in conveying and deriving meaning, namely the distinction between (effectively) encoded meaning, on the one hand, and all other contingent meaning, on the other (Levinson 1995: 222; cf. introduction to this book). This view of meaning forces us to examine signs in their usage across a wide variety of contexts. In order to verify hypotheses about the semantic invariant of a linguistic expression (Wierzbicka 1996), one **must** consider the interaction between context and encoded meaning. The rich details of context, which constitute input for implicatures, are as important in the analysis as the encoded semantics (Hanks 2005).³ While it is necessary to separate context-independent meaning (encoded semantics) from context-dependent meaning (pragmatics), these aspects of the full repertoire of meaning-conveying mechanisms used by speakers are complementary and interdependent, and therefore must be studied simultaneously.

2.2 Deixis

Deictic expressions depend inherently on context for interpretation of reference (Lyons 1968: 275ff., Fillmore 1997 [1971], Clark *et al.* 1983, Hanks 1990). Nevertheless, as Bühler (1982 [1934]) established, deictic expressions have a context-free dimension, namely the idea of a stable **deictic centre** or **origo**, which anchors simple notions ‘I’, ‘here’, and ‘now’. To be absolutely clear, it is the idea of an origo which is stable across contexts, not the actual coordinates of the origo on a particular occasion of use. (Further, demonstratives

³ That one can believe in encoded meaning and maintain this view has been missed by some critics of the semantics–pragmatics distinction. Consider, for example, the claim that adoption of a semantics–pragmatics distinction implies a belief that semantics is ‘much more philosophically important than pragmatics’ (Lakoff 1987: 171). Note, however, that while Langacker (1987: 154) once rejected the pragmatics/semantics distinction, he has more recently written of a language ‘code’, referring to meaning that is ‘schematic at the level of established linguistic convention’, which then takes on ‘a specific value when instantiated in a particular instance of language use’ (Langacker 1997: 234). It is encouraging to see recognition within the cognitive linguistics tradition that information acquired and/or derived on the basis of a sign in context is not necessarily generalizable as part of the value of the sign itself. For further discussion of these points, see Grice (1989), Levinson (1995: 210, 1997), Wilkins and Hill (1995: 212–214, 252–253).

will also encode other context-independent features such as number – singular, plural etc. – gender, distance, among many other things; see below, pp. 29–67.) Bühler pointed out (as others have since done; for example, Wierzbicka 1996) that the core conceptual elements of the *origo* are not semantically decomposable.⁴ Being challenged to define them will either ‘induce the language theorist into esoteric philosophical abysses or to respectful silence’ (Bühler 1982 [1934]: 13). Accordingly, we may treat ‘I’, ‘here’, and ‘now’ as semantically primitive (Wierzbicka 1996; cf. 1972: 16). In addition to Bühler’s core deictic domains of person, place, and time, there is a further stable and simple deictic notion, namely a primitive demonstrative meaning, a symbolic indicating function which I gloss here as *DEM*.⁵ This meaning cannot be reduced or expressed in terms of some other feature of the *origo*, and is simply an essential indicating expression, rooted ultimately in the meaning of the prelinguistic pointing gesture (Halliday and Hasan 1976: 57; Liszkowski 2006, Tomasello 2006 – see Chapters 2–4 of this book). It is abstract with respect to an exophoric/endophoric distinction (Himmelmann 1996, *pace* Halliday and Hasan 1976: 59, Diessel 1999: 93ff.). It is useful for drawing attention, but attention-direction is not a semantically specified function. Its use presupposes that an addressee can know what it is referring to. The following section elaborates on the nature of the *DEM* function, which is at the semantic core of any demonstrative.

2.2.1 *Demonstratives*

Demonstratives are often associated with indexicality in physical space, but they do not necessarily encode spatial meanings (Kirsner 1979, 1993, Hanks 1990, 2005). They can be used in a range of domains. Himmelmann (1996) establishes four major functions of demonstratives cross-linguistically: situational, discourse, anaphoric, and recognitional (cf. Lakoff 1974, Diessel 1999). The situational usage, involving ‘reference to entities in the surrounding situation’, has also been referred to as exophoric (Halliday and Hasan 1976: 33, Diessel 1999: 6) and has traditionally been treated as primary

⁴ This nondecomposability is at the level of intensional meaning and is criterial of the status of a semantic meaning as primitive (Goddard 1998). The global meaning of demonstratives in their situated usage can be complex, as examples described in this chapter illustrate. My argument is that this global meaning is context-enriched and is not equal to the stable schematic encoded core meaning of the expression.

⁵ This technical term is mnemonic for **demonstrative**. It could also be glossed by the semantically most general demonstrative in a language. I choose not to use an English demonstrative to gloss it, since there is disagreement about which English term is the unmarked one (compare Halliday and Hasan 1976: 59 who claim *that* as basic with Wierzbicka 1980: 27 and Dixon 2003: 81 who claim *this* as basic in English).

with respect to other, endophoric functions such as discourse and recognitional deixis. Diessel, like others before him, claims that the exophoric use of demonstratives is ‘the basic use from which all other uses derive’ (1999: 93). Diessel’s arguments, however, appealing to grammaticalization, acquisition, and markedness, do not establish basicness at the level of synchronic adult semantics. I follow Kirsner (1979), Hanks (1990), and Himmelmann (1996: 223, 242) in maintaining a monosemy bias and regarding the simplest demonstrative meanings as general with respect to exophoric versus endophoric reference. The exophoric use of a semantically general demonstrative is one domain of application (i.e. part of the term’s extension), not a distinct semantic meaning (i.e. not specified at the intensional level). This is possible because a referent is always a conceptual entity, irrespective of whether it is instantiated in the linguistic discourse or (also) the situational context.

The basic function of demonstratives is not to specify **where** something is, but rather to specify **which one** you are talking about (see Fillmore 1982: 43–44). I cannot rely on a demonstrative to provide the content for a reply to a where question. Thus, if I ask *Where is my copy of War and Peace?*, you cannot answer by saying *This book* or *This one*.

When I use a demonstrative – for example, by asking *Have you read this book?*, with reference to *War and Peace*, sitting with other books on a table between us – you and I have a coordination problem (Schelling 1960, Lewis 1969, Clark *et al.* 1983). We need to identify one and the same book, and we need to rely on each other to do so in the situation. My use of the word *this* alone is not enough for our problem to be solved. Its core meaning ‘DEM’ announces that I assume we can converge on a solution to the problem, given information accessible to both of us. Speakers and listeners both assume a kind of conversational maxim which Clark *et al.* (1983) call the principle of optimal design: ‘The speaker designs his utterance in such a way that he has good reason to believe that the addressees can readily and uniquely compute what he meant on the basis of the utterance along with the rest of their common ground’ (Clark *et al.* 1983: 246). That a speaker may intentionally be vague or misleading does not refute the idea of such a principle. The presupposition referred to here functions as a maxim, not as a rule. It is a ‘presumptive heuristic’, which ‘informs (rather than strictly governs) conversation’ (Levinson 2001). The key to understanding this point is that the maxim is not so much a rule for the **speaker** to follow than it is for the **addressee** to **presume** that the speaker follows. This means that speakers are not assumed to intentionally invite inferences that would require information unavailable to their addressees. Accordingly, addressees assume that speakers will have tailored their inference-inviting utterances to make reference only to information known to be in common ground (that is, information not just shared between speaker and addressee, but known by each to be so shared;

D’Andrade 1987: 113, Clark 1996: 92ff., Enfield 2000: 45).⁶ Suppose that by saying *this book* I mean ‘the copy of *War and Peace* sitting on the table here now’. You may figure this out as a result of some mutually obvious perceptual salience of the book, of our current shared focus of attention on the book, of my physically positioning the book in your field of attention (by holding it up, for example), or of my directing your attention to the book (perhaps by pointing at it). In another situation, the book may not be perceptually accessible at all, but nonetheless salient enough (e.g. in the current focus of conceptual attention in the discourse) to provide us with a mutually unique solution. The basic demonstrative **meaning** does not encode the idea of which book ‘this book’ is (in this case, *War and Peace*). But the effect of its use and the application of interpretive heuristics in a textured context gives rise to a sensible composite utterance, that is a piece of speech connected to a piece of the world.

All languages have at least one demonstrative determiner, and there is almost always a second term, often a third, and sometimes more (Anderson and Keenan 1985, Diessel 1999, Dixon 2003). This adds further dimensions to the process of interpretation of a single demonstrative described by Clark *et al.* (1983). A semantic distinction encoded in a second demonstrative term provides the addressee with an additional clue for solving the coordination problem of reference. It seems that the most common meaning added by a second term is spatial. When a semantically more specific demonstrative comes into paradigmatic opposition with a semantically more general one, important system-internal pragmatic effects arise, and these are not discussed in standard typological accounts of demonstratives (Lyons 1977, Anderson and Keenan 1985, Diessel 1999). While standard analyses depict pairs like *this* and *that* as symmetrically expressing the notions ‘near speaker’ and ‘far from speaker’, respectively, the symmetry is emergent rather than coded. Instead, the two terms may form an informativeness scale (Horn 1989, Atlas 2005, Levinson 2000: 79), whereby the use of a semantically less specific or weaker form (given that a semantically more specific or stronger form is an option in the same grammatical context) implies the converse of the stronger form, yet without semantically encoding it. Thus, as I argue below, since *nan4* **encodes** ‘not here’, *nii4* often **implies** that the referent is something ‘here’. But the two terms show different degrees of semantic specificity.

⁶ Note, however, work by Keysar *et al.* (2000) arguing that addressees occasionally ‘consider some referents from their own perspective, even when they know that these referents are inaccessible to the speaker’ (Keysar *et al.* 2000: 37; cf. Barr and Keysar 2004). It is not clear, however, whether this has consequences for the idea of the ‘optimal design principle’ as I apply it here.

Within the realm of spatial meaning, as specified for example by a semantically more specific demonstrative, it is crucial to distinguish a semantic specification of location from one of distance. Semantic content can specify **where** something is without specifying **how far away** it is (*pace* Diessel 1999: ch. 3 *passim*).

While *nii4* and *nan4* can serve both exophoric and endophoric functions, discussion here is restricted to exophoric uses only. The data are from video recordings of spontaneous interaction among Lao speakers in situations that involve reference to, and manipulation of, real objects in interactional space. This method of investigation into demonstratives is necessary in order to remedy the problems of standard analyses. In complement to these naturally occurring examples, I have also discussed these scenes with consultants and have sometimes asked for judgments as to whether another demonstrative could have been used. Further, I applied an elicitation technique for probing demonstrative distinctions by setting up situations depicted in a set of drawings (Wilkins 1999), covering a range of hypothetical scenarios designed to vary with respect to proximity of a concrete referent – such as a book – to speaker and addressee, at various distances and in various conditions (see van Geenhoven and Warner 1999: 56). The technique works especially well when the situation elicits a judgment that only one demonstrative is possible. In most cases, however, speakers report that either demonstrative could conceivably be used (although one may be preferred). The problem here is that on any given occasion, speakers do not ‘use either’ – they must and do select one. The choice is not arbitrary, but is a function of the semantics of the alternative demonstratives, in combination with rich contextual factors.⁷ Speakers’ a priori intuitions are that *nii4* is used for things near the speaker, while *nan4* is used for things far from the speaker, but observational data show that referents of ‘proximal’ *nii4* are often quite far away, indeed as far or further than many referents of ‘distal’ *nan4*. Conversely, things marked by distal *nan4* may be physically very close to the speaker. The same is true for English *this* and *that*.

A central point in understanding how demonstratives are used in exophoric functions is that human interaction transforms merely physical space into meaningful space. This cannot be measured by observing purely physical spatial properties of interactional scenes (Hanks 1990). Interlocutors monitor and model each other’s mental states (Goody 1995, Clark 1996), and this includes their construals of physical space as controlled, possessed, shared,

⁷ And when it appears that the conditions are such that either of the demonstratives would lead equally well to the right unique solution to the current coordination problem, then it still matters which one you use: the choice will reflect or create interpretations of spatial or social cohesion (Hanks 1990; cf. Enfield and Stivers 2007).

separated, and so on, by one another (Goodwin 2000a). To gain insight into this phenomenon, it is necessary to look at real examples of demonstrative use in all their richness.

2.2.2 *The extensional range of 'here'*

This analysis assumes a basic spatial meaning 'here'. While I can use *here* to refer to places of seemingly arbitrary extension – the Earth, the city of Nijmegen, Office 275, a place on the bottom right corner of my front tooth – the idea of 'here' remains conceptually discrete. The intension of 'here' is a place, prototypically 'where I am now'. When I say *John's here* I tell you where he is, but not how far away. Physical measurements of specific context-situated spatial extensions of the word *here* would tell us little about its intension. Instead, it is necessary to consider properties of the speech location not as mere space but as interactional space, partly in terms of physical attributes of the space itself (including spatial arrays which speakers form with their bodies and other props as they interact; Schefflen 1976), but most importantly, in terms of phenomena affecting the interpretation of space emerging from the dynamics of the human interaction itself (Goffman 1963, Kendon 1977).

People interact in places like rooms, buses, fields, gardens, and shops. There are physical boundaries within these places (doorways, walls, vegetation), which do not absolutely prevent communication across them, but which may nevertheless be regarded as social boundaries, and may be treated as if they were true physical barriers to communication (Goffman 1963: 151ff.). People physically orient themselves in space to create interactional closure (Kendon 1977: ch. 5; see also Schefflen 1976, Goodwin 2000a, among many others) and, thus, effective boundaries to communication. These spatial closures emerge from, and are defined by, joint communicative activity and are located in the minds and manners of people rather than in the physical setting itself. Logically independent of such closures, but often coinciding with them, is a conceptually defined area that may be called the here-space. The here-space is not necessarily defined with reference to physical facts, but is simply the place or area which one considers as 'here' at a certain moment for a certain purpose. Directly analogous to the puzzle of how interlocutors use demonstrative reference to solve coordination problems of reference is the puzzle of how interlocutors coordinate their convergent understanding of the perimeter of 'here' at a given moment.

Just one of the factors that can define a speaker's here-space is his or her engagement area, the place which is, at a given moment, the conceived site of a person's currently dominant manual and attentional engagement. Kendon describes something similar in what he calls the transactional segment: The

‘space extending in front of a person ... which he is currently using in whatever his current activity may be’, ‘the space into which he looks and speaks, into which he reaches to handle objects’, which he will ‘endeavor to maintain’ (Kendon 1977: 181). The transactional segments of two interlocutors may ‘overlap to create a joint transactional space’ (ibid.). The factors that determine the perimeter of a person’s engagement area during interaction are many and varied, and constantly competing. The perimeter of one’s engagement area shifts constantly, from moment to moment, like a fluid and invisible shell encircling specific configurations of people and things in space. Suppose that I am one of two people standing talking at a cocktail party. We face each other, creating a tight engagement area. If we are joined by a third, we broaden our focus, and our encircling shell reshapes (as we step back to include the new interlocutor), then re-cohering (or being reconstituted), now larger, around the three of us as we form a fresh symmetrical array (Kendon 1977: ch. 5). When I approach a market table and enter into discussion with a merchant – who is facing me on the other side of the table – about the price of an item lying on the table between us, an engagement area is created, embracing the two of us and our mutual focus of engagement, the item to be purchased. When the sale is made and the merchant turns to wrap the item, her visual and manual preoccupation with that task causes the previous two-person engagement area to fade, and her adjusted focus creates a new solo-engagement area, now much smaller, momentarily excluding me and the table with goods on it (see below, pp. 38–61). I am assuming that no more than one engagement area can be dominant at any one moment, although of course there can be a number semi-active or implicit. Thus, the engagement area bridging the market table is soon resumed after the merchant is finished with her momentary solo-engagement in wrapping the purchased item.

In interactional spaces, such as the open market places in which many of the examples described in this chapter were collected, people continually monitor the dynamic activation of engagement areas and the interactional borders they create. The perimeters of these areas are **felt** by people, and the monitoring of such perimeters is seldom if ever explicit. Sensitivity to these socially meaningful spaces is suggestive of the inexplicit, unarticulated yet ‘cultivated disposition, inscribed in the body schema’ described by Bourdieu (1977: 15; cf. Hanks 1990, 2005). This conception of interactionally established perimeter interacts with referent location, resulting in a number of logical possibilities of relevance to the selection of demonstratives, to be examined below. An object (e.g. a book) that a speaker wants to refer to with a demonstrative may or may not be in his or her own engagement area at the relevant moment. Also, the engagement areas of speaker and addressee may or may not overlap at the relevant moment. The engagement area perimeter is just one possible determinant of the more abstract here-space perimeter, and

Table 2.1 *Logical possibilities for coincidence of location of referent, speaker's here-space, and addressee's here-space.*

Is referent in here-space of speaker?	Is referent in here-space of addressee?	
a. yes	yes	entails $HERE_{SPKR} \cap HERE_{ADDR}$
b. no	no	i. where $HERE_{SPKR} \cap HERE_{ADDR}$ ii. where $HERE_{SPKR} \neq HERE_{ADDR}$
c. yes	no	entails $HERE_{SPKR} \neq HERE_{ADDR}$
d. no	yes	entails $HERE_{SPKR} \neq HERE_{ADDR}$

as I argue below, it is the here-space perimeter that is directly relevant for demonstrative selection. (Sharing an engagement area at a given moment does not entail sharing a here-space.) The logical possibilities for coincidence of location of a referent, speaker's here-space, and addressee's here-space are shown in Table 2.1.

There are a number of interactional reasons for speakers to make the construals they do, when they do. For instance, one's choice of demonstrative can signal a claim about whether one's own here and the here of one's addressee are shared. This can be influenced by strongly suggestive aspects of the physical situation, by subtleties of the social situation (e.g. distinctions arising out of politeness – it may be impolite to assume one's space to be shared with that of one's addressee), or by deliberate manipulation. But ultimately the choice is made by the speaker with the intention that the demonstrative selected be the one most likely to result in the coordination problem of reference being solved by the addressee's convergence with the speaker on the right referent. Let us now turn to the Lao system of demonstratives, in particular the two demonstrative determiners *nii4* and *nan4*.

2.3 Lao demonstrative determiners *nii4* and *nan4*

Lao has five demonstrative elements, none of which may appear independently as agent or undergoer noun phrases (cf. Enfield 2007: 99–100). Only two – *nii4* and *nan4* – are genuine demonstrative determiners, since only these two can generally be used as nominal modifiers in simple noun phrases like 'this book'. This puts *nii4* and *nan4* in paradigmatic opposition to each other. The remaining three are demonstrative adverbs (*han5* 'there', *phii4* 'here', and *phun4* 'yonder'). The two nominal-modifying demonstrative determiners *nii4* and *nan4* have different degrees of semantic specificity. *Nii4* is the more semantically general member of the pair, encoding nothing more than a basic primitive demonstrative meaning (paraphrasable as 'this (one)',

‘the (one) that is mutually salient enough in this context for you to know which one I must mean’, expressed here as DEM). *Nan4* is the semantically more specific member of the pair, encoding the same basic demonstrative meaning plus the specification that the referent is something ‘not here’.

- (1) a. *nii4* = DEM
 b. *nan4* = DEM NOT HERE

These two create an informativeness scale (Levinson 2000: 79), whereby use of the semantically more general of the two can implicate (but not entail) the converse of the semantically more specific of the two. The semantically more specific *nan4* encodes that the referent is located in a region that is ‘not here’, while the semantically more general *nii4* (given that *nan4* is an option in the same grammatical environment) can suggest the converse, namely that the referent is ‘here’.

This kind of analysis of multi-term demonstrative systems – as consisting of one semantically general term that may be enriched by inference in opposition to other, semantically more specific, terms – was prefigured by Halliday and Hasan (1976: 59) and by Wierzbicka (1980: 27), and in collaborative research by the Space Group at the MPI, Nijmegen (Wilkins 1999, van Geenhoven and Warner 1999: 58). The idea of characterizing the distinctive semantics of a marked distal form as ‘not here’ was earlier suggested by Wierzbicka (1980: 27). To pursue ‘not here’ as the operative semantic distinction in the Lao case allows preservation of a discrete intensional semantic value invariant across all uses, while remaining consistent with wide variation in context-specific extensional value, due to often highly complex external factors. While it may be tempting to suggest that *nii4* entails ‘here’, I pursue the claim that it does not, for three reasons. First, it is not necessary to do so in order to account for the data, and thus is favoured by Ockham’s principle of simplicity. Second, refraining from attributing any spatial information to the intension of *nii4* allows a monosemy analysis of *nii4* such that its endophoric uses (not explored in this chapter) are handled under the same single meaning as its exophoric uses. This is also favoured by Ockham’s principle of simplicity. Third, according to informants’ judgments, *nii4* is always a conceivable possibility in contexts where *nan4* has been used but the reverse is not true. This is evidence that the two forms are not symmetrical in semantic content, and more specifically, that *nii4* is semantically more general than *nan4*.

We turn now to examples that provide evidence for the claims that (a) of the two Lao demonstrative determiners, *nii4* is semantically general and *nan4* is semantically more specific; (b) the added semantic content of *nan4* specifies that the referent is something ‘not here’; (c) neither *nii4* nor *nan4* encode semantic distinctions of ‘distance’; and (d) the choice of demonstrative

determiner is made with reference to the perimeter of the speaker's here-space as conceived in the interlocutors' common ground. These conceived perimeters are emergent from factors of the interaction, including active engagement areas, physical features of the interactional space, and assumptions about addressees' access to relevant information for inference.

2.4 Uses of semantically more specific *nan4*

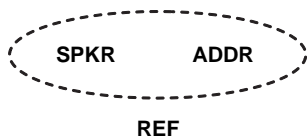
The message common to all exophoric uses of *nan4* is that the referent is something 'not here'. This section describes a range of circumstances that can lead a speaker to decide that sending this message, rather than simply using the semantically general demonstrative *nii4*, will be more likely to result in a convergent solution by speaker and addressee to the coordination problem at hand. Following are examples of two situations in which a referent may come to be conceived of as 'not here'. In the first, the speaker has a saliently bordered here-space, where the referent is plainly not in it. In the second, the addressee has a salient here-space perimeter, where the referent is plainly in it and the speaker does not share here-space with the addressee. This entails that the referent is, for the speaker, 'not here'.

2.4.1 *Situation: speaker has salient here-space perimeter, referent is not in it*

CASE 1: Speaker's here-space perimeter is defined by engagement with the addressee; Referent is not in it.

One situation that leads to construal of a referent as being 'not here' is when it is outside the speaker's here-space perimeter because of a salient here-space closure arising from an engagement involving both speaker and addressee ((bi) in Table 2.1).⁸

(2)



In Figure 2.1, a number of children are in a boat and are about to cross the river. The boat is about thirty metres from the camera. Immediately to the left of the camera (1–2 metres away) are two people engaged in conversation.

⁸ The here-space perimeter is represented here, and below, as a dotted line circle or ellipse. The abbreviations Spkr, Ref, and Addr stand for 'speaker', 'referent', and 'addressee', respectively.

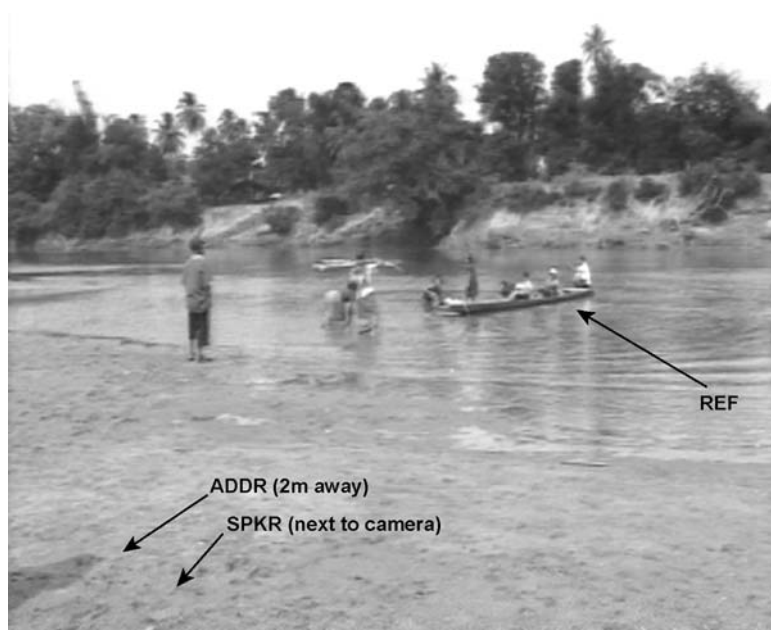


Figure 2.1 ‘There’s only children in *nan4* boat.’

One speaker (a woman) notices the children in the boat getting ready to cross, and remarks ‘There’s only children in *nan4* boat.’

The speaker’s use of *nan4* to refer to the boat signals that the boat is ‘not here’ from her point of view at that moment. Her current dominant engagement is a conversation that does not extend to or involve the people visible in Figure 2.1, who are out of earshot and engaged in the riverside action. It is clear to the interlocutors that the referent is in a place well outside of the here-space defined by their current local conversation. Thus, the speaker can assume that her interlocutor will also understand the boat as being in a place ‘not here’ at the relevant moment. (An example from moments later in this same scene, discussed below (cf. Figure 2.8), discusses a very different construal, by a different speaker, of the same physical speech setting.)

Figure 2.2 shows a customer and a merchant in the middle of a sale. In Figure 2.2a, they together create a tight engagement area with joint attention on physically coordinating an exchange of produce for cash.

In Figure 2.2b, as the merchant has now taken the money, and the two withdraw, the merchant breaks out of the established engagement space by turning to her right and pointing, with gaze and outstretched index finger, to some tomatoes, asking ‘*Nan4* ones, won’t you put some in?’



Figure 2.2a Customer (right) and merchant (left) with joint attention and coordinated manual action of exchanging goods for cash.



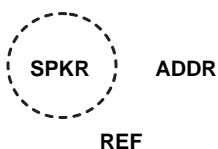
Figure 2.2b '*Nan4* ones, won't you put some in?'

The merchant already knows what the customer is planning to cook, knows that the recipe requires tomatoes, and knows that the customer hasn't purchased tomatoes yet. At the moment that she wants to draw the addressee's attention to the tomatoes, a common here-space is already mutually salient to the two interlocutors (the space defined by the tight engagement area arising from their joint attention and manual engagement in synchronizing the cash exchange illustrated in Figure 2.2a). The tomatoes are out of the salient perimeter of that here-space. Thus, the speaker uses the demonstrative *nan4* in referring to them, specifying that they are 'not here'. The status of the referent as not here is determined by facts about how perimeters of interactional space acquire mutual salience for interlocutors. The perimeter of 'here' at this moment cannot be determined with reference to geometry alone (Hanks 1990).

CASE 2: Speaker's here-space perimeter defined by tight solo-engagement, excluding addressee; Referent is not in it.

A referent may be construed as not here when a salient perimeter of the speaker's here-space arises from the speaker's solo-engagement, and the referent is beyond that perimeter. This case is like the previous in that the referent is out of the speaker's here-space perimeter, but it is unlike the previous in that the here-spaces of the speaker and addressee do not overlap ((bii) in Table 2.1).

(3)



In Figure 2.3, both speaker (merchant, second from left) and addressee (customer, far left) are very close to the referent. Figure 2.3a shows the merchant standing next to the customer at the corner of her market table. However, the merchant is engaged in a verbal exchange with the man at the back of the stall (who has not participated in any interaction involving the customer). She is also solo-engaged in the manual task of wrapping up the customer's purchase of meat. (This example also has an element in common with 2 above, in that the here-space of the speaker, which does not contain the referent, is partly determined by the speaker's interaction with another person. In this case, however, that person is not the addressee.)

At this point, the customer asks the price of a piece of meat on the corner of the table immediately between her and the merchant. The merchant turns to



Figure 2.3a Merchant (second from left) in an engagement area whose perimeter excludes customer (far left).



Figure 2.3b ‘*Nan4* one is 16 per kilo.’

look down at the meat in question, still engaged in wrapping the previous purchase, and answers ‘*Nan4* one is 16 per kilo.’⁹

This example shows a referent physically highly proximal to the speaker (literally, within arm’s reach) being referred to by a so-called distal demonstrative. I checked with consultants as to which demonstrative would be expected, giving only information about the spatial layout of speaker, referent, and addressee, and no information about engagements or other interactional dynamics. All consultants preferred *nii4*. Proximity of a referent to a speaker suggests default inclusion in here-space, but this is overridden in the example described here by a salient interactionally established here-space perimeter very close to the speaker. How could a model of demonstratives as semantically encoding distance deal with the paradox of a distal demonstrative being used in an indisputably proximal context? The example shows that we cannot appeal to objectively measurable distinctions of physical distance, but must acknowledge that what determines the choice here is a conceived perimeter in the space, again not as mere space but as interactional space.

CASE 3: Closure of speaker’s here-space perimeter determined by physical properties of the interactional space; referent is not in it.

While here-spaces are conceptual entities established by people’s **interpretation** of space, physical properties of a situational setting can have a strong effect on such interpretations. Physical entities like walls, doorways, conventional wall-projecting items like chair backs and other obstacles are conventionally acknowledged barriers that speakers and addressees can mutually recognize (Goffman 1963, Schefflen 1976). Such physical properties are common ground for interlocutors based on perceptual evidence (Clark *et al.* 1983: 247, Clark 1996: 112); that is, they are common ground by virtue of being plain for both to see. Conceived barriers of this kind can override other features of common ground as determinants of here-space perimeters.

I am arguing that a speaker’s physical or perceptual access to a referent can be a factor in selection of one of the two Lao demonstratives, but I am not saying that visibility or access are encoded in their semantics. The speaker’s lack of access to a referent is what causes it to be conceived of as ‘not here’, and it is the conception of the referent as not here, rather than the lack of access, that is encoded in the demonstrative *nan4*. The descriptive linguist thus has reason to show caution before claiming that a factor such as visibility is semantically encoded in a demonstrative system (see e.g. Anderson and Keenan 1985: 290, Diessel 1999: 40–42). This is not to deny the possibility that visibility might be encoded in a demonstrative system, but to suggest care

⁹ The currency unit is the Kip (approximately 8500 = USD1 at the time of recording). References to prices 16 and 17 are in thousands.

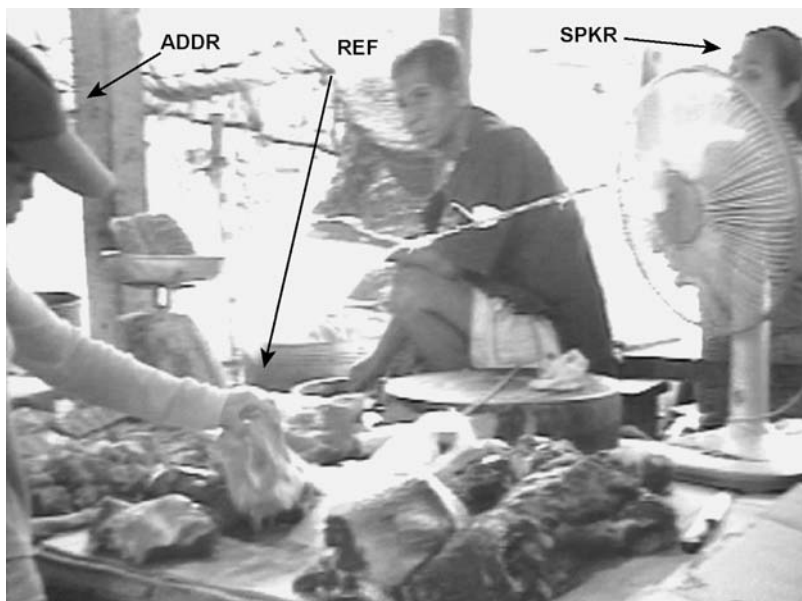


Figure 2.4 ‘*Nan4* one is 17 per kilo.’

in distinguishing whether a feature such as visibility is part of encoded meaning or is a contextual factor that forces the choice of another, more general semantic meaning.¹⁰ I describe below a case in which the **addressee’s** access to the situation can also, via quite a different mechanism, determine a speaker’s choice of demonstrative.

We now consider two examples in which a physical barrier in the interactional space contributes to recognition of a here-space perimeter (and, accordingly, selection of an appropriate demonstrative).

In Figure 2.4, a merchant (to far right) is behind her market table, obscured by an electric fan on one side of the table. There is also a man sitting behind the table (on another table) to the other side, and a plastic bag on a string attached to the fan (to deter flies and wasps) hovering in the air and blocking the view across the table (left centre).

Each of these factors, as well as the vacant space between the merchant and the table itself (since she is standing away from it), form a perceptually salient physical barrier in the common ground of the interlocutors and is thus naturally construed as a here-space perimeter. Further, the merchant is in a

¹⁰ The point made here may also shed light on how a demonstrative term with semantic encoding for visibility of a referent might grammaticalize.



Figure 2.5a Joint attention on water inlet to washing machine. ‘And what did you do about water?’

solo-engagement, wrapping the customer’s purchase. The customer asks the price of a piece of meat on the table (using *nii4*, for reasons outlined in section 2.5, below), and the merchant replies, from her position behind the electric fan, ‘*Nan4* one is 17 per kilo’. The merchant’s choice to encode the referent as not here is due to a here-space perimeter that has arisen partly from the presence of physical barriers in the interactional space and partly from her own self-enclosing solo-engagement at the moment of speech.

I have described to consultants the spatial geometry of speaker, addressee, and referent in this example, and they are unanimous that the speaker could have used proximal *nii4*. But given the interactional dynamics in this real context, the speaker had good reason to use the more semantically specific form *nan4*. Once again, information about physical distances between interlocutors and a referent alone cannot account for the distribution of demonstrative selection.

The three parts of Figure 2.5 show two women – let us call them the hat woman and the older woman – engaged in a discussion about how to operate a second-hand washing machine that the older woman has just acquired. The machine was previously at the house of the hat woman, and the older woman is asking the hat woman details about how to operate it. Figure 2.5a shows the

two women's joint attention on the water inlet of the washing machine, as the older woman points at it with her index finger, asking 'And what did you do about water?', referring to the arrangement connecting the water inlet to a water supply.

Up to this moment, the women have created a tight engagement area encompassing the two of them and the washing machine, where their joint attention has been focused for some time. They have been physically huddled over the washing machine. Then, as illustrated in Figure 2.5b, the hat woman looks up and out of the current engagement area, to a tap on the wall, which is connected to the washing machine's inlet hose. She points at the tap with gaze and outstretched index finger, saying 'Keep (it) on the whole time, *nan4* thing.'

Although the referent is less than two metres away (as shown in Figure 2.5c), there is nonetheless a significant physical barrier – the washing machine together with the body of the older woman – between it and the speaker, and the presence of this barrier is perceptually salient to both speaker and addressee. The referent is well and truly out of the speaker's physical reach. Note that the speaker does not have access to the referent via the right side of the washing machine (the side closest to the camera). The video recording is taken from outside a window looking in, and the washing machine is against a wall.

This example demonstrates how a barrier to physical access can contribute to establishment of a relevant here-space perimeter. There is also a contribution from the rather restricted perimeter of the engagement area the speaker had already set up with the older woman (illustrated in Figure 2.5a). Note that this is maintained by the older woman's visual attention on the speaker and not on the referent in Figure 2.5b.

2.4.2 *Situation: addressee has salient here-space perimeter, referent is in it, speaker is not.*

Suppose an addressee has a saliently bounded here-space perimeter, arising from his or her engagement with someone other than the speaker or with some object. If the referent is in that space (especially if it is the focus of the addressee's attentional or manual engagement), and if the speaker is outside that space, then the referent must be 'not here' for the speaker (regardless of the positively defined extent of the speaker's here-space). A speaker will thus use *nan4* to encode the idea that the referent is not here. Martina Faller (personal communication) suggests that it is problematic for the speaker's here-space not to be positively defined, thus begging the question of how the location of 'not here' can be computed. Let me clarify the logic of what I am claiming: Premise 1 – the referent is in the here-space of the addressee; Premise 2 – the here-space of the speaker does not overlap with that of the addressee; Conclusion – the referent must be 'not here' for the speaker. I am



Figure 2.5b 'Keep it on the whole time, *nan4* thing.'



Figure 2.5c Referent of deictic gesture and demonstrative in Figure 2.5b: tap connection for water inlet to washing machine.

suggesting that as long as the addressee's here-space perimeter is saliently and clearly defined and there are salient and clear reasons for the speaker and addressee to be regarded as not sharing respective here-spaces, then there is no need to know the positive extension of the speaker's here-space in order to calculate the location or identity of the referent on the basis of a speaker's 'not here' specification. Schematically:

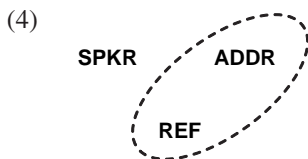


Figure 2.6 shows a customer (the hat woman), accompanied by a child, standing at a market table.

The woman is busy purchasing a number of different types and quantities of meat (visible on the table). There is an established active engagement area for



Figure 2.6 '*Nan4* insects!' (Referents – wasps – at table's edge in front of hat woman's hand; speaker behind camera to right.)

her that includes the array of meats on the table and the merchant, who is on the other side of the table (out of camera frame, to right). There are a number of other people standing around, conversing and eating together, in front of the adjacent market table (away approximately three metres to the right of the hat woman and the child in Figure 2.6), and these people form their own, quite separate, interactional closure. One man in the group of people to the right (out of camera frame) notices that the child in Figure 2.6 is trying to play with some wasps that are buzzing around the meat on the table, and warns the hat woman, implying that she should intervene. In Figure 2.6, she is backing away from the table, looking at the man as he speaks. He uses the demonstrative *nan4* in referring to the wasps (i.e. as *mèng2 qanø-nan4* [insect MC.INAN-*nan4*] ‘*nan4* insects’), specifying that they are ‘not here’ (**for him**).

This use of *nan4* accords best with the likely shared assumption about perimeters of speakers’ and addressees’ respective here-spaces, and the location of the referent with respect to these spaces. It is plain to both speaker and addressee that the referents – wasps – are firmly in the addressee’s here-space, the perimeter of which coincides with the interactional closure of the purchase-at-market-table scenario. The man standing away to the right (out of shot) has up to this point had no interaction with the addressee, and thus is not part of the engagement which had established the addressee’s here-space closure. It is as if out of politeness he should use *nan4* here, to signal explicitly that he does not presume that he shares a single here-space with his addressee (a stranger). Further, he is within an engagement area of his own, in an inward-oriented circular formation he has created with his conversational interlocutors (Kendon 1977: ch. 5). Thus, the referent is for the speaker not here, and this is encoded by his use of *nan4*. This solution is sensitive to shared assumptions and likely counter-assumptions about common ground in the context, since the active here-space closures are mutually salient to all concerned. The choice could not be explained without taking the interactional context into account (e.g. by appealing to a distinction of physical distance).

I asked consultants to judge which demonstrative would be appropriate, describing only the physical spacing of speaker, referent, and addressee in this example and suppressing information about the interactional context. As in other examples discussed in this chapter, consultants said the speaker could have used either *nii4* or *nan4* for that referent. But in the actual situation the speaker chose *nan4*. Given the richer interactional context, consultants agreed that *nii4* would have been very odd, and probably confusing. Again, interactional dynamics are crucial in selection of demonstratives, as a means for gluing word to world in composite utterances. An analysis of demonstrative distinctions which cannot take these factors into account would be too blunt a tool.



Figure 2.7 ‘*Nan4* ones are 2500.’

Figure 2.7 shows the hat woman to the left, buying vegetables at a stall attended by an older woman (‘older merchant’, centre) and a younger woman (‘younger merchant’, to right).

The hat woman and the older merchant are in a tight engagement area, created by their joint attention on selecting various vegetables at the leftmost end of the market table where they are standing. They are physically oriented to each other, both bending over, jointly attending to the produce on the table between them, forming a closure that excludes the younger merchant, who is standing well back and observing. There is an interactional (and also visual) gravity pulling the two women together to the exclusion of the third, and since this is perceptually apparent to each of the three women in this scene, it is common ground between them.

The hat woman asks the price of some tomatoes (using *nii4*, for reasons outlined in section 2.5) that are in the centre of the tight engagement area maintained with the older merchant. It is the younger merchant behind who responds, saying ‘*Nan4* ones are 2500 [per kilo].’

The addressee and referent are together in a tightly established here-space whose perimeter does not include the speaker (the younger merchant, to right). The speaker’s use of *nan4*, asserting that the referent is something not here, indexes her non-sharing of that here-space, where the referent is located. Recall the logic: Premise 1 – referent is in addressee’s here-space; Premise 2 – speaker does not share addressee’s here-space; Conclusion – referent cannot be in speaker’s here-space, regardless of the positively defined extension of that space.

As mentioned in discussion of the previous example, the mere spatial layout of speaker, addressee, and referent is not sufficient for consultants to agree that one or another demonstrative is obligatory or especially preferred. In this case, the fact that the addressee is already engaged in interaction with a third person (who is neither speaker nor addressee) is crucial for the speaker's selection between the two demonstrative determiners. It is only when we understand the complex dynamics of the interaction in real time that the interactional logic of the appropriate choice of demonstrative becomes clear.

2.4.3 Summary

I have shown in this section that *nan4* encodes an intensional meaning consisting of DEM – which specifies that the reference of what it marks should be recoverable by the addressee with reference to salience in currently relevant common ground – plus a specification of the spatial location of the referent, namely that it is not here. This is revealed by its use in two types of situation. In the first type, the perimeter of the speaker's here-space is salient in the interaction and the referent is not within that perimeter. This situation can arise in several ways, depending on factors that determine a publicly salient closure of the speaker's here-space (including perceived spatial closure created by interactional engagement, and more tangible spatial closure created by distance or physical barriers). In the second situation, the perimeter of the addressee's here-space is salient in the interaction and the referent is within that perimeter, and the speaker's here-space does not overlap with that of the addressee.

2.5 Uses of semantically more general *nii4*

What all uses of *nii4* have in common is that the speaker is not saying that the referent is something not here. Given that Lao speakers have one other choice of demonstrative determiner (*nan4*), which encodes that the referent is not here, the use of *nii4* often implies (but never entails) that the referent is here. In other words, *nii4* and *nan4* form an informativeness scale (Levinson 2000: 79), with *nan4* the strong member of the pair. While *nan4* is genuinely specified for awayness (a locational specification that must not be confused with marking for distance), the status of *nii4* as proximal arises entirely by inference due to an oppositional relation in the system to a semantically more specific alternative. *Nan4* and *nii4* are not equivalent in semantic complexity, and their status as distal and proximal, respectively, is qualitatively distinct. *Nan4* has a semantic specification for where the referent is, while *nii4* does not. Consider now two main situations in which *nii4* is used in spontaneous interaction.

2.5.1 *Situation: no here-space perimeter is particularly salient*

In the following examples, the speaker's engagement area is large and diffuse, such that there is no particular contextually salient perimeter of here within the local interactional space. As a result, it becomes inappropriate to specify as 'not here' even a referent quite far from the speaker. Such cases support the view that *nii4* is nonspecified with respect to place (or hereness). The fact that no particular here-space perimeter is mutually apparent to interlocutors means that an explicit specification of the referent's location with reference to here (or not here) is information the addressee cannot use. Speakers are presumed by addressees to take addressees into account in formulating utterances which are intended to be interpreted by those addressees (Sacks and Schegloff 1979/2007, Clark *et al.* 1983, Clark 1996). If there is a mutually obvious here-space perimeter in an interaction, then an addressee can expect a speaker to refer to it in inviting a certain interpretation. But when no such perimeter is salient in the common ground, then a speaker cannot expect reference to the notion of a here-space perimeter to be of use in interpreting the speaker's utterance. Instead, an addressee would struggle to comprehend the relevance of such a specification. I offer two examples.

In Figure 2.8, a man is descending a river bank, calling out to children in a boat, who are setting off to cross the river. The shot is taken from the same location as Figure 2.1, swiveled about 120 degrees to the left, only seconds after the moment captured in Figure 2.1.

The speaker wants the children in the boat to cross the river and embark at a particular landing, and he calls out '*Nii4* boat-landing! Cross to *nii4* boat-landing!', index finger pointing with outstretched arm. The referent (namely the landing) is at least 100 metres away. Figure 2.1 shows both the boat and the landing (which can be seen as a bare patch on the river bank to the centre left).

While the referent is surely far from the speaker, there is no reason in this context why the speaker would want or need to specify that the landing he means is 'not here'. He is walking down the river bank alone, in no interactional or similar engagement that may create a salient locally closed here-space. He calls out to the children in the boat, who are over thirty metres away, and his voice is loud enough to be heard by everyone around. At this point, he is broadcasting his speech across a wide space, encompassing the entire open area of the river and its two banks, including all of those who can hear him as he shouts. There is no particularly salient here-space perimeter in the common ground, and he thus uses the semantically general demonstrative *nii4* for this referent.

It is worthwhile explicitly comparing this example with Figure 2.3b, above, in which the supposedly distal form *nan4* was used with reference to



Figure 2.8 ‘*Nii4* boat landing!’

something within hand’s reach of the speaker. In the Figure 2.8 example, the supposedly proximal form *nii4* refers to something more than 100 metres away from the speaker. Distance cannot be what distinguishes the meanings of these two demonstratives.

A similar situation pertains to Figure 2.9. Figure 2.9a shows a woman, index finger pointing, arm outstretched. She is pointing across an irrigation ditch to one of a pair of buffaloes being led along by her brother-in-law (illustrated in Figure 2.9b).

The woman is in the middle of a typically diffuse exchange of conversation with a number of people around the irrigation ditch and its banks, at varying distances. Such diffuse conversation is a common mode of interaction in a Lao village setting, partly due to the physical layout of Lao living areas (thin walls, windows without panes, doorways without doors, open-air verandas). People freely and comfortably converse with others who are not in the same spatial enclosure, and anyone within earshot of an interaction may normally assume involvement (unlike cultures such as that of, say, urban United States; Goffman 1963). The conversation just before the speaker’s pointing gesture



Figure 2.9a (‘Which one’s pregnant?’) ‘*Nii4* one.’



Figure 2.9b Referent of *nii4* in 2.9a is the buffalo in front.

in Figure 2.9a involved people some twenty metres to the woman's left, on the same side of the irrigation ditch, as well as the man leading the buffaloes along on the other side. One person has commented on the size of one of the buffaloes, which it turns out is pregnant. At this point the interactional engagement area is broad and diffuse, encompassing people (and buffaloes) on both sides of the ditch, with speakers' contribution being broadcast more than directed at anyone in particular. In this larger interactional context there is no salient local closure around the speaker at the point when I (holding the camera) ask her 'Which is the pregnant one?' Her response is to point, with index finger and gaze, saying '*nii4* one', as illustrated in Figure 2.9a. Figure 2.9b, from the same spot just seconds later (with the camera swiveled about ninety degrees to the right), shows that the referent, the furthest buffalo, is quite far from the speaker, at least twenty-five metres.

The diffuse, broadcast nature of the interaction and the speaker's lack of any salient current engagement or attention at a more local level than the passing buffaloes and the scattered individuals within earshot means that there is no impetus for her to encode the referent (the pregnant buffalo) as 'not here'. She shares with her interlocutors no particularly salient idea of a here-space perimeter between her and the buffaloes, and so none is referred to.

2.5.2 *Situation: speaker has salient here-space perimeter, referent is in it*

In many situations, there is a salient here-space perimeter around the speaker, and the referent is located within it (corresponding to (c) in Table 2.1). The speaker cannot expect any addressee to have assumed that the referent would be construed as not here, so the selection of *nan4* – encoding not here – would be pragmatically confusing. This point makes reference to a principle of audience design (cf. Sacks and Schegloff 1979/2007), by which addressees can assume that messages have been formulated with them in mind. If a here-space perimeter is salient common ground in an interaction, then a speaker knows that an addressee will assume that the obvious presence of that here-space perimeter will have been taken into account by the speaker in formulating that utterance for that addressee. Thus, when the most publicly obvious construal of a referent's location is that it is 'here' for a given speaker, then, in order to avoid being misconstrued, that speaker should not encode it as 'not here', and therefore must use the semantically general demonstrative *nii4*.

CASE 1: Speaker's here-space perimeter includes referent, excludes addressee. The clearest examples are where the speaker is actually holding or otherwise manipulating an object.

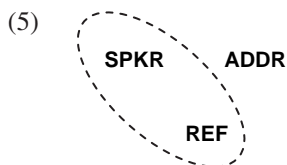


Figure 2.10 ‘*Nii4* one, should I slice in pieces?’

Figure 2.10 shows the merchant from previous examples (Figures 2.3, 2.4, and 2.6) sitting on her market bench at her cutting block. The merchant has placed on the block a piece of meat that the customer has selected, and is now going to ask the customer whether the meat is to be sliced into pieces. She index-finger points to the meat, with her fingertip almost touching it, saying ‘*Nii4* one, should I slice in pieces?’

Informants are unanimous that the speaker could not have used *nan4* here. Not only is the referent in the spatial centre of a currently dominant engagement area, but it is the very focus of the dominant engagement itself. Since this is obvious to both speaker and addressee, the speaker has no reason



Figure 2.11 ‘You like *nii4* one?’

to encode the referent as not here. Note that this is not due to the objectively measurable spatial proximity of the referent to the speaker. In a different interactional situation, an object just as close to the speaker could indeed be encoded, by *nan4*, as not here (see Figure 2.3).

CASE 2: Speaker’s here-space perimeter includes referent and addressee. Often a referent is the focus of the speaker’s and hearer’s joint attention, thus a central part of what defines a shared engagement area, and, in turn, a shared here-space.

(6)

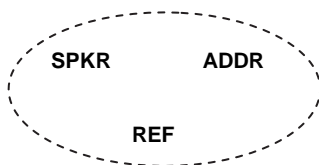


Figure 2.11 shows two adults and a small child in a classic joint attention scene (Tomasello 1999). The child has a set of plastic animal figures, which she has set out on the floor between the three. As the man looks on, the woman asks the child ‘Which one do you like?’ The child indicates the one she likes by touching it. The woman then points to the same figure, asking ‘You like *nii4* one?’



Figure 2.12 ‘*Nii4* row, *nii4* row, and that’ll be enough.’

The engagement area these interlocutors have created has the referent at its physical and attentional centre. The specification ‘not here’ is inapplicable, and so the semantically general demonstrative *nii4* is used.

Figure 2.12 shows a man and a woman harvesting lettuce from a market garden. They are each holding a knife, and the woman is carrying harvested lettuce, while the man is doing the cutting. They have not yet harvested enough for current requirements, and at this moment are jointly attending to the garden bed, considering which rows of lettuce to cut next. The woman says, using the knife as a pointer, ‘*Nii4* row, *nii4* row, and that’ll be enough’, as the man motions to grab the lettuce plants and begin cutting.

This joint attentional scene puts the man, the woman, and the referent (the rows of lettuce) in one and the same engagement area, where the interlocutors share attentional focus on the referent, and, in turn, share one and the same here-space. Again, a specification of ‘not here’ would be inapplicable, and so the semantically general demonstrative *nii4* is used.

2.5.3 Summary

Nii4 encodes an intensional meaning consisting only of the simple demonstrative meaning DEM (specifying that the reference of what it marks is expected by the speaker to be recoverable by the addressee, given current common ground). It lacks any specification of the spatial location of the referent. But it often picks up an inferred meaning ‘here’ by paradigmatic opposition to the available more semantically specific term *nan4*, which encodes ‘not here’. This is revealed by the use of *nii4* in two main types of situation. In the first type of situation, there is no salient perimeter of the speaker’s here-space in the interaction (for example when the interaction is spatially diffuse), rendering an explicit reference to ‘not here’ irrelevant to the referential coordination task at hand. In the second type of situation, the perimeter of the speaker’s here-space is salient in the interaction and the referent is saliently within that perimeter (with subtypes in which the addressee is included, and not included, respectively, in the speaker’s here-space). This situation rules out the possibility of encoding the referent (by *nan4*) as ‘not here’, leaving the semantically general term *nii4* the only possibility.

2.6 Addressee location, speaker–addressee dynamics, and common ground

Some demonstrative systems (e.g. systems with three or more elements) are reported to be person-oriented, where the spatial location of the addressee is a reference point for the location of the referent (Diessel 1999: 50, cf. p. 39, Anderson and Keenan 1985: 284). I have argued that the Lao system is, in semantic terms, speaker oriented, in that the distinction between the two demonstratives is anchored in a speaker origo. I have also shown, however, that addressees contribute to spatial and interactional dynamics, and can thereby play a crucial role in influencing speakers’ construal of the relevant spaces, and in turn determine which demonstrative is used.¹¹ Another relevant factor affecting speakers’ choices of demonstrative is not just their own construal of interactional space, but, more importantly, their model of how addressees expect messages will be tailored for them. I have mentioned already that speakers frame their linguistic choices under the assumption of a

¹¹ Özyürek (2000) shows that addressee location has an effect on the nature of shared interactional space, and, in turn, on representational gestures. A number of authors argue forcefully for consideration of the addressee in our most general understanding of how communication works (Sacks and Schegloff 1979/2007, Goodwin 1979, 2006, Clark 1996, Stivers *et al.* 2007).

maxim of audience design (cf. Sacks and Schegloff 1979/2007). In all the examples described here, whether it is an interactional engagement area, some physical obstacle, or some culturally conventional barrier which determines a here-space perimeter, what matters is speakers' confidence that addressees will be able to model the speaker's intended construal of here-space – that is, the construal an addressee believes he or she is 'intended to infer on the basis of common ground' (Clark *et al.* 1983: 257). Speakers tailor their utterances so that addressees are not required to make reference to information that the speaker does not know or assume that they have access to. In turn, addressees expect speakers' utterances to be tailored so as not to depend on information that is not assumed by speakers to be already shared with addressees. In the kinds of interactional situations examined here, the physical location of an addressee can determine the availability (i.e. by visual or other perceptual access) of certain information, and in turn, speaker–addressee common ground, and in turn, a speaker's choice of demonstrative determiner. Thus, addressee location plays a crucial role in the selection of demonstratives, not only due to addressees' part in affecting the status of shared space (e.g. by physically or interactionally contributing to the establishment of speakers' here-space perimeters), but also due to their part in determining how speakers' messages are designed. And none of this entails any reference to the addressee in the semantics.¹²

The following example is illustrative of the pragmatic effect that speaker location can have on demonstrative selection. In Figure 2.13a, the hat woman has picked up a cabbage in her hand, and asks the younger merchant (see Figure 2.7, above) the price, saying '*Nii4* one, how much?' By holding it up, she positions the cabbage for the attention of the younger merchant (Clark 2003).

The addressee, standing behind the market stall, does not know the price. The older merchant, who normally runs the stall, is at this moment some fifteen metres away from her stall (behind and to the right of the younger merchant), engaged in conversation with another stall-owner in a neighbouring stall. The younger merchant turns back to her right, calling to the older merchant 'How much is *nii4* one?' (Figure 2.13b).¹³ This is the utterance of interest here.

¹² This is not to say that addressees cannot be referred to in the semantics of other demonstrative systems. Özyürek and Kita (2001; cf. Kita and Dickey 1998: 65–66) argue that the Turkish demonstrative *su*, traditionally referred to as encoding medial distance (in opposition to proximal *bu* and distal *o*), in fact encodes the non-attention of the addressee to the referent. Thus, a referent of *su* is 'something you (the addressee) are not attending to now'.

¹³ Note that the pillar on the left of Figure 2.13a is the same pillar as the one to the right of Figure 2.13b. This gives an idea of the distance between speaker and referent (about two metres).



Figure 2.13a ‘*Nii4* one, how much?’



Figure 2.13b ‘How much is *nii4* one?’

This example is important because it allows only the semantically more general term *nii4*. When consultants are shown this scene, they are unanimously adamant that the speaker in Figure 2.13b could not have used distal *nan4* when addressing the older woman, standing far from the market stall pictured. At the same time, consultants are agreed that if the addressee had

been the hat woman, holding the referent in her hand, then the speaker could have used either *nan4* or *nii4*. This may be contrasted with Figure 2.4, in which the speaker uses distal *nan4* in a situation almost identical to this one in terms of the purely spatial geometry of speaker and referent placement.

If the addressee in Figure 2.13b had been the hat woman, and if the speaker had used distal *nan4*, this would have suggested a construal of the hat woman's here-space as being closed off to the speaker, and therefore that the cabbage (undeniably in the hat woman's here-space) was, for the speaker, not here. As shown in examples such as Figures 2.4, 2.6, and 2.7, this would not be unusual. But given the interactional and physical geometry of the speaker and addressee combination in Figure 2.13b, consultants find it inconceivable that the speaker would have encoded the cabbage as 'not here'. It would be bad audience design. The speaker has not at this point been interactionally engaged in any way with the older merchant. Thus, speaker and addressee do not currently share any common ground concerning currently active spatial closure. The most salient closure that the speaker and addressee in Figure 2.13b can assume to be common ground is the perimeter provided by default in the context, namely the conventionally defined closures separating each individual market stall from the next. In this context, the market stall perimeters are default candidates for here-space perimeters. The notion of speaker's 'here' that the speaker can expect to be most saliently available to the addressee is thus 'here at this market stall'.

This is a case of interlocutors jointly solving a coordination problem with a Schelling choice based on assumptions about common ground (Schelling 1960: 54). Even though specific interactional factors could have contributed to the establishment of a locally divisive here-space perimeter **within** the market stall area (dividing the speaker's space from the space of the woman holding the referent and thus putting the referent in a place 'not here' for the speaker), the addressee in the case of Figure 2.13b (the older woman, out of sight) could have had no knowledge of those factors, since she was away from the scene. The speaker knew this, and the older woman – as addressee – knew that the speaker knew this. Accordingly, the addressee expected the speaker to have taken this common ground into account in designing the utterance (Clark *et al.* 1983: 257). This claim is supported by consultants' unanimous rejection of distal *nan4* as a possibility for the scene in Figure 2.13b.

2.7 Appropriating the here-space perimeter for contrast between two referents

In many languages, different demonstratives may be used to express contrast between two possible referents (e.g. English *this one*, *not that one*). Often,

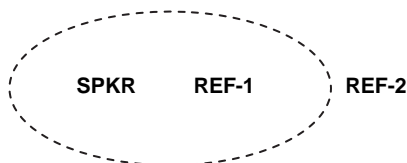
a Lao speaker will use a single demonstrative to refer to two things contrastively (where contrast is achieved nonverbally, for example, by finger-pointing). There are cases, however, in which *nii4* and *nan4* are used contrastively. The basic principle is to arbitrarily place the here-space perimeter between the two objects, as long as the objects are viewed on an away-going axis.

2.7.1 *Recruiting the outward-radiating border between here and not here*

Lao speakers often use *nii4* and *nan4* to distinguish contrastively between two objects that are placed at different distances on a line radiating out from the speaker. Figures 2.14a–d show a sequence in which two women are discussing two newly purchased hand towels, which are the focus of their joint attention. The speaker, on the right, indicates contrasts between the two towels, first in terms of price, and second in terms of thickness. As she speaks, she points with her index finger to the relevant towel, in the sequence shown by Figures 2.14a–d.

The speaker makes contrastive reference to the towels by arbitrarily setting the perimeter of here-space as bisecting the outward-radiating line from one towel to the next, construing the further towel as not here (by using *nan4*), and using the unmarked *nii4* to refer to the remaining towel (the nearer one). This strategy is common when a pair of objects are in this kind of alignment: Nearer to and further from the speaker, respectively, on a line radiating out from the deictic centre. This is illustrated schematically in (7).

(7)



This is a clear case in which relative distance does determine the selection of demonstratives, yet without being encoded semantically. As discussed in section 2.2.2, the reference of ‘here’ is intensionally discrete yet extensionally elastic. Once a perimeter of here-space is set, something outside it will naturally be ‘far’ in comparison to something inside it. It is perhaps because near and far are scalar notions that people have regarded them as the operative distinction in demonstrative semantics (see Kemmerer 1999).

When objects are arrayed crosswise, not on that outwards-radiating line, then *nii4* and *nan4* are generally not used contrastively, as the example in

Figure 2.14a ‘*Nii4* one cost 4000.’Figure 2.14b ‘*Nan4* one cost 7000.’Figure 2.14c ‘*Nan4* one is thicker . . .’



Figure 2.14d ‘... than *nii4* one.’



Figure 2.15 Fourteen uses of *nii4* (none of *nan4*) in a twenty-second sequence of reference to different control buttons.

Figure 2.15 shows. In Figure 2.15, two women discuss how to operate a washing machine (cf. Figure 2.5, above). In a twenty-second sequence, fourteen demonstrative references – all *nii4* – are made to various buttons, as the two women maintain joint attention on the array of buttons, continually

touch-pointing throughout. They are discussing which buttons are to be pressed at which times during the washing machine's cycle.

In this case the speakers are unable to use the contrastive strategy seen in the previous example, whereby the here-space perimeter is deliberately set to bisect an outward-radiating line from self to one object and to the next one.

2.7.2 *Using distal nan4 as 'the other one' where a binary choice is salient*

There are contexts in which *nii4* and *nan4* are used contrastively when the objects in question are not in a line radiating away from the speaker. If two referents to be contrasted are a salient pair of parts of a thing (identical in some relevant sense), one may be referred to by *nan4*, with a reading translatable into English as 'the other one'. The remaining referent of the two is referred to, by contrast, with *nii4*. I observed an instance in which a small child was trying to get up onto a tractor cart, and she was having trouble doing so, because she was trying to get on at the wrong end. The child's mother was sitting, watching, about five metres away, with the cart sitting crosswise, that is, with the line from one end of the cart to the other exactly perpendicular to the line radiating out toward the cart from the mother's deictic centre. Neither end of the cart was closer to the mother than the other. Nevertheless, she said to the child 'No, get on at *nan4* end!', meaning 'No, get on at the other end.' One explanation is that the speaker is transposing her deictic centre to that of the child (Bühler 1982 [1934], Haviland 1993). There are two ends to the cart, and one is closer to the addressee than the other. It is the one 'not here' from the addressee's point of view which is referred to by distal *nan4*.

2.8 Concluding remark

I summarize the findings of this chapter as follows. Of the two Lao demonstrative determiners, neither encodes information about distance, and only one encodes information about location. Neither *nii4* nor *nan4* tells the addressee 'how far away' a referent is, and only *nan4* says 'where' it is. Pragmatic inference alone gives rise to the association of *nii4* with things 'near speaker' or 'here', and *nan4* with things 'far from speaker'. What is shared by the two forms is a core primitive demonstrative meaning, a specification that a speaker is referring to something that he or she assumes is salient enough in the common ground to be uniquely identifiable by an addressee as the referent the speaker is referring to. The two demonstrative determiners interact system-internally, forming an informativeness scale,

with one weak or semantically general member, and one strong or semantically specific member, the latter being more informative. Their semantic meanings are abstract and general but nevertheless stable and specifiable. Hand-in-hand with this, there are pragmatic factors that can be systematically described.

This study shows that factors such as distance of referent from speaker, location of addressee with respect to referent, and visibility of referent to speaker or addressee can each play a role in the selection and deployment of demonstratives, yet without needing to be specified in their semantics. Lao speakers' choices between the two available demonstratives are influenced by conceived extensions of here-space. These conceptions are determined, in turn, on the basis of pragmatic factors emergent in the dynamic interactional situation, including physical barriers, perimeters of engagement area created by interlocutors' manual and attentional focus, and conceived gravitational pull between interlocutors and objects, among other things. In addition, given the omnipresence of an interactional maxim of audience design, speakers also take into account their addressees' presumed access to information relevant for inferring the extension of the speaker's conceived here-space. In this way, demonstratives in all languages are both speaker-anchored and addressee-anchored. To what extent the contributions of these contextual and interactional factors are culturally variable remains to be seen, but considerations of this kind need to be included in further cross-linguistic research on the semantics of demonstratives.

Standard linguistic typologies of demonstratives would lead one to expect a two-term system of demonstrative determiners such as Lao *nii4* and *nan4* to encode a simple spatial distinction of proximal versus distal. But the data described in this chapter show that such an analysis lacks important insights. For instance, it would not tell us why the 'proximal' demonstrative can refer to things both closer and further away than things which may be referred to by the 'distal' term. A standard analysis both overestimates and ill-defines the semantic content of these expressions, and fails to acknowledge sufficient richness in the contribution of context and pragmatic implicature to demonstrative usage. The advantage of the approach taken here over an analysis based on a symmetrical opposition of distance-marking is that it shows how a stable and discrete intensional value can map onto greatly variable analogue extensional values. This cannot be understood without acknowledging a distinction between discrete, conventionally encoded intensional meaning and other aspects of meaning arising in and from context. Semantics and pragmatics are distinct, but they must be treated as equal parts in a whole analysis of composite utterances.

In conclusion, let us consider demonstratives not as a linguistic phenomenon but as an instance of a more general class of device for building

composite utterances – the symbolic indexical. As we shall see by taking this chapter together with the following two, demonstratives are no different from hand-pointing gestures in several key respects. They provide a universally present yet locally conventionalized solution to the problem of attaching conventional signs in spoken utterances to physical objects and other token states of affairs. They constitute small, closed classes of distinct types which tell interpreters to go looking in the context for some referent to be glued to the proposition supplied by the conventional sign components of the utterance. Their closed class status means they contrast with each other in conventional semantic terms, and so beyond saying ‘go look for a referent’, they help both to narrow the search space and to express some kind of stance towards that referent.

3 Lip-pointing

Bust a move.

Young MC, 1989

Lip-pointing is a widespread form of deictic gesture, a systematic and conventionalized behaviour attested in at least Southeast Asia, the Americas, Africa, Oceania, and Australia. ‘Lip-pointing’ is not an ideal label, since there is more involved in these gestures than merely pointing with the lips. A salient aspect of the gesture is the action of making one or both lips protrude as if using one’s lips to point to or at a referent. But there is almost always an accompanying quick raising of the head and chin, and orientation of eye gaze towards the referent, with an occasional eyebrow-raise in addition. The lip-pointing itself is part of a set of deictic actions involving the head and face, and is to be regarded as part of the larger deictic system in which it occurs. Within the relatively little existing research on pointing in general (Hewes 1981, Haviland 1993, Kita 2003, *inter alia*), lip-pointing receives little mention. Apart from Sherzer’s (1973) description of the ‘pointed-lip gesture’ of the Kuna people of San Blas in Panama (cf. also Sherzer 1983, 1993), little can be found on the topic, other than in the context of discussion of hand gesture (e.g. Key 1962: 94, Hewes 1981: 265, Poyatos 1983: 114–116, Wilkins 2003: 174–179) or in linguistic descriptions of demonstratives and other deictic forms (e.g. Feldman 1986: 196).

This chapter describes lip-pointing behaviour of Lao speakers, with additional discussion of lip-pointing more generally (with reference to available data from elsewhere). The data are from a set of over thirty examples of lip-pointing observed in video recordings of informal interaction in rural and semi-urban villages of lowland Laos.

3.1 Form

The term ‘lip-pointing’ should not be taken to mean that only the lips are involved. None of the Lao examples involve the lips alone. Additional actions of chin-raise/head-lift, gaze direction, and eyebrow raise are usually involved. The combination of all these could be termed a full-blown lip-point, as in Figure 3.1.



Figure 3.1 Full-blown lip-point, involving chin, head, gaze, eyebrows.

These various articulatory components are technically separable, but in the full-blown lip-point they conspire to point most emphatically. I am mostly concerned in this chapter with the lips, although I also discuss associated kinesic actions involving head, face, and gaze, as well as forms of hand-pointing. I claim a special role for the function of eye gaze in lip-pointing and in related types of deictic gesture involving the head and face.

3.1.1 *The lips*

There are six logical possibilities for two lips involved in a deictic gesture, varying on the parameters of (a) whether the lips are parted versus together, and (b) whether both lips or just one (upper versus lower) is the primary articulator.

Of these possibilities, all except 4 are attested in the limited available cross-cultural data.

Figures 3.2 and 3.3 are examples typical of Lao speakers, showing the lips parted, involving upper and dual-lip protrusion, respectively (i.e. instantiating cells 1 and 2 of Table 3.1).

Table 3.1 *Six logical possibilities in lip-pointing form.*

	lips parted	lips together
<i>upper lip primary</i>	1	4
<i>both lips</i>	2	5
<i>lower lip primary</i>	3	6



Figure 3.2 Lao speaker, lips parted, upper lip primary protrusion.

The following two examples show equal lip-protrusion, with lips parted, and together, respectively (Figures 3.4 and 3.5, instantiating cells 2 and 5 of Table 3.1). These are from recordings of a speaker of Jahai (Aslian,



Figure 3.3 Lao speaker, lips parted, equal lip protrusion.

Mon-Khmer) in peninsular Malaysia, collected by Niclas Burenhult (cf. Burenhult 2000), and kindly made available for inclusion here.

The next example illustrates a Kuna speaker lip-pointing, with lips together, protruding equally (Figure 3.6, instantiating cell 5 of Table 3.1; Sherzer 1973: 117).¹

Finally, the illustration in Figure 3.7 of a Colombian lip-pointing shows protrusion of the lower lip (instantiating cell 6 of Table 3.1; Saitz and Cervenka 1972: 33).

Sergio Meira reports (in personal communication) that Brazilian Portuguese speakers lip-point with a protruding lower-lip, more or less as illustrated in Figure 3.7, while indigenous groups such as speakers of Tiriyo from Northwestern Amazonia use a very different looking parted-lips equal-protrusion gesture, more or less as illustrated by the Lao example in Figure 3.3.

¹ Figure from Joel Sherzer, 1973, 'Verbal and nonverbal deixis: the pointed lip gesture among the San Blas Cuna', *Language and society* (Cambridge University Press) 2, 117–131, reproduced with permission.



Figure 3.4 Jahai speaker, lips parted, equal protrusion. (Image courtesy of Niclas Burenhult.)

In mentions of lip-pointing in the literature where no illustration is supplied, it is difficult to tell which of the formal possibilities in Table 3.1 are realized. Descriptions like ‘puckering the lips’ suggest equal lip protrusion (with lips either parted or together), but it is hard to be sure. Key (1962) says the following about lip-pointing among indigenous Bolivians:

Many Indian tribes point with their lips; we recorded the Movima, Tacana, and the Ayoreo as using this gesture. But there were differences in executing this gesture; the Movimas do not accompany it with a thrust of the head as the other tribes do, but simply protrude their lips to point out an object. (Key 1962: 94)

Lip-pointing among Navajo speakers is described as follows: ‘purse the lips together as if you were to kiss someone, keeping the lips pursed move the head in the direction of the object you wish to refer to’ (Larry



Figure 3.5 Jahai speaker, lips together, equal protrusion. (Image courtesy of Niclas Burenhult.)

DiLucchio, the Navajo Nation, personal correspondence; cf. DiLucchio 1998/1999²).

Wilkins describes lip-pointing among the Arrernte in Central Australia as ‘made by orienting one’s head face-on towards a referent while protruding both lips (sometimes just the bottom lip)’ (Wilkins 2003: 187). Wilkins also cites Eckert and Hudson’s description of lip-pointing among neighbouring speakers of Pitjantjatjara: ‘extend the bottom lip and raise the chin at the same time’ (Eckert and Hudson 1988: 87). Neither of these descriptions indicate whether the lips are usually together or apart, but clearly the lower lip is often the one primarily protruding (i.e. instantiating cells 3 or 6 of Table 3.1).

3.1.2 *Movements of the head, chin, and eyebrows, and the direction of gaze in association with lip-pointing*

A common pattern is for the lip-pointing gesture to be preceded or accompanied by orientation of the head towards the target (so that the lips will point

² Thanks to David Wilkins for directing me to this source.



Figure 3.6 Kuna speaker, lips together, equal protrusion.

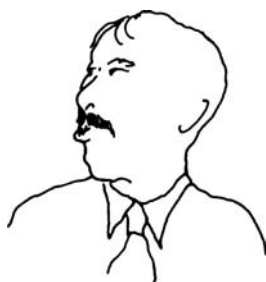


Figure 3.7 Colombian Spanish speaker, lower lip primary protrusion.

in the right direction). In Figure 3.8a, the speaker is asked ‘Where is your house?’ She responds in Figure 3.8b by turning her head towards the referent (in view, approximately ten metres away) and making the lip-point gesture (without speech).

In the next example (Figure 3.9a), the speaker (whose hands and attention are occupied with looking at a photograph album) is asked ‘Where did you get married?’ She responds in Figure 3.9b by turning her head towards the

referent and making a lip-point, saying ‘At home, over there in H. K. Village’ (a village about one kilometre away, in sight).

In the example in Figure 3.10, the speaker (on the left) does not laterally re-orient her head in order to lip-point, since she is already facing in the right direction. In Figure 3.10a, the woman on the right asks ‘Where is Grandma and the others?’ The speaker replies (Figure 3.10b) by raising her head in combination with a lip-point, saying *phun4* ‘Over there’.



Figure 3.8 (a) ‘Where is your house?’; (b) Head orientation with lip-pointing.



Figure 3.9a ‘Where did you get married?’



Figure 3.9b ‘At home, over there in H. K. Village.’



Figure 3.10 (a) ‘Where is Grandma and the others?’; (b) ‘Over there.’

The next example also demonstrates marked contrast between the poses before, and during, the gestures involving the lip-point. In Figure 3.11a, the speaker (at right, partly obscured) is suggesting to his colleague (at left) that they move to a cooler spot some ten metres to the right of the camera. The colleague asks ‘Where?’ and the speaker replies in Figure 3.11b with the emphatic particle *ndêê4* ‘Here!’ (a contraction of *nii4 dêê4* DEM FAC. FILLIN), turning the head to orient the face in the direction of the referent location, raising the head and chin (the latter visibly jutting out), orienting gaze to the referent location, and raising the eyebrows, opening the eyes wide.



Figure 3.11 (a) 'Where?'; (b) 'Here!'

In all these examples, the speaker's gaze is aligned with the lip-point. That is, whenever someone lip-points in a certain direction, they are also looking in that direction. This is further discussed, below.

A further finding concerning gaze is that lip-pointing is apparently restricted to cases when the addressee is looking at the speaker. In the example in Figure 3.12 (from the same sequence as the previous example), the speaker (at right, obscured) is suggesting to the addressee a better place to sit. At first, the addressee (at left) is not looking at the speaker (in Figure 3.12a), yet as the speaker speaks he index-finger points. It is only when the addressee turns and looks at the speaker that the speaker then performs the lip-point (in Figure 3.12b).

Further, as is also common throughout the Lao data, as soon as the lip-point in Figure 3.12b has been performed, the addressee then immediately directs his own gaze towards the referent (i.e. to where the speaker's gaze was directed).

Indeed, in all of the more than thirty lip-pointing examples in the Lao data, as far as it can be ascertained, the speaker makes the gesture only while the addressee is looking at him. By contrast, many kinds of hand-pointing gesture are made while the addressee is not looking directly at the speaker. This may



Figure 3.12 Lip-pointing when addressee is looking.

relate to the greater visibility of hand-pointing gestures than lip-pointing gestures in peripheral vision.³ However, while in peripheral vision one is perhaps more likely in the case of hand-pointing to see that one's interlocutor is pointing, one presumably would be no more clear about exactly where or at what the speaker is pointing. Accuracy is then a second problem. Regarding accuracy, some have claimed that lip-pointing is less precise than hand-pointing (especially index-finger pointing; Hewes 1981: 265). This is doubtful (Wilkins 2003). First, if lip-pointing is a switch for the deictic vector of gaze, then it inherits the high accuracy of gaze as a pointer (cf. section 3.2.3). Second, in any case, the accuracy of a pointing gesture is, as Adam Kendon puts it, more 'an achievement of the interaction' than a computation from geometric information (cf. ch. 2, above, and Hanks 1990).

3.1.3 *Sequence of speech, lip, and hand-pointing*

The Lao data show that lip-pointing occurs in structured combination with speech, and with other gestures such as various forms of hand-pointing.

³ Thanks to Adam Kendon for pointing this out.

The possibility of speech occurring together with lip-pointing is naturally better facilitated when speakers conventionally use a lips-parted form (as they tend to do in Laos). Formal observation of relative sequence of hand- and lip-pointing reveals that all logical sequences are possible: lip-point followed by speech and/or hand-point, speech and/or hand-point followed by lip-point, and lip-point and speech and/or hand-point concurrently. Following sections discuss functions of some of these different sequence types.

3.2 Function

3.2.1 *Distribution of lip-pointing in terms of semiotic function: focus on location/identity*

In the data discussed in this chapter, lip-pointing only occurs in location-focus utterances (cf. Chapter 4), that is, acts of direct ostension in which the location or identity of a referent in the physical environment is in focus. By saying that something is in focus, I mean that it constitutes the part of a proposition that makes an assertion beyond what is presupposed by that proposition (Lambrecht 1994: ch. 5). For example, in *John took your newspaper* (with stress on ‘John’ and not on any other constituent), the presupposition is that ‘someone took your newspaper’, the assertion is that ‘that someone was John’, and the focus is *John*. ‘WH-questions’ usually presuppose everything but the question word. For example, *Who took my newspaper?* presupposes ‘someone took my newspaper’. Or, *Where did he put it?* presupposes ‘He put it somewhere.’ Thus, minimal answers to ‘where?’ questions are focal. This is the status of lip-points in the Lao data. All of the more than thirty examples arise as answers to ‘Where?’ questions or as otherwise explicit focal assertions of location or identity (e.g. ‘It’s there’, ‘It’s this one’).

Consider the examples discussed above. The lip-point in Figure 3.2 is in immediate response to a question as to where a certain kind of wood may be found in the local forest. Figure 3.3 shows an immediate response to the question ‘Where is your house?’ The lip-pointing gestures in Figures 3.9–3.12, above, and 3.14–3.16, below, show the same pattern, being answers to questions like ‘Where did you get married?’, ‘Where are Grandma and the others?’, and ‘Where is your old TV?’ (see discussion of the individual examples for details). In each of these cases, by virtue of being an answer to a ‘where?’ or ‘which one?’ question, the location or identity of the referent is in focus in the composite utterance which incorporates the lip-point.



Figure 3.13 Index-finger pointing to a non-focal referent.

In none of the examples observed is a lip-point made when the referent location or object is not focal in the utterance. This narrow restriction on semiotic function is in distinct contrast with forms of hand-pointing, which may occur in a broader range of contexts, and, in particular, in contexts in which the gesture indicates something non-focal in the communicative exchange (see Chapter 4). Let us now consider an example.

A common type of deictic gesture in Lao discourse involves the index finger (or the thumb, if the referent is directly behind the speaker), where the arm is not fully extended, neither torso nor gaze are turned in the direction of the referent, and the gesture is brief, casual, relatively close to the body, and usually not held over time (see Chapter 4 for elaboration). An example comes from an informal interview in which the investigator (at far left in Figure 3.13) asks the man second from left, ‘What grade do your two daughters study in?’, and the man replies by saying ‘Fourth and fifth grades, here in our village’, making the index-finger pointing gesture as he speaks.

Here, the pointing gesture is not held for any duration, and while eye gaze happens to be aligned with the gesture, it was not turned to match the direction of the pointing gesture. The physical target of the gesture is the village school (some 500 metres away and out of sight), yet the focus of the spoken component of this composite utterance is ‘fourth and fifth grades’. The location and identity of the school that the daughters go to constitutes non-focal information. In such contexts, hand-pointing is normal, while lip-pointing is not attested.



Figure 3.14 (a) ‘Where is the old set?’; (b) ‘(It’s) up in the house’; (c) Index-finger pointing following lip-pointing.

3.2.2 Coordination of lip-pointing with forms of hand-pointing

Close sequential combinations of lip-pointing with other deictic gestures are common. Some generalizations can be made as to how lip-pointing combines with other kinesic actions.

First, when a lip-point and some kind of hand-point (usually an extended index-finger point) at the same referent occur in the same utterance, as they do in nearly half the examples in the Lao data, it is more often the case that a lip-point is followed by an index-finger point. This is a kind of up-grading, as if the speaker, having lip-pointed, feels that this is somehow not enough, and then re-phrases by finger-pointing.⁴ It may be that the lip-point has a more ‘recognitional’ flavour (Sacks and Schegloff 1979/2007), indicating a message from the speaker to the addressee along the lines of ‘I think you know where/which one I mean’, i.e. that the thing or location being pointed out is presumed to be shared knowledge at some level (Chapter 2 above; cf. Stivers *et al.* 2007).⁵ If so, then having lip-pointed, a speaker may feel that this expressed assumption of shared knowledge is perhaps too presumptuous, and this is then corrected by a more neutrally informative index-finger point. Enchronic contingencies of the interaction are what generate these patterns. We now consider four examples which provide empirical evidence for this analysis.

In the following example (Figure 3.15), the speaker, who has just bought a new television set, is asked ‘Where is the old set?’ The set is in a room almost directly above him (he is sitting downstairs). He first replies (in Figure 3.14b) by lip-pointing with raised eyebrows and head tilted far back, gaze directed to the location of the referent, as he says *juu1 theng2 huan2* [be.at above house] ‘(It’s) up in the house’. The speech and lip-pointing are simultaneous. He

⁴ I thank Steve Levinson for helping me develop this idea.

⁵ This would accord with lip-pointing being considered more informal and familiar (see section 3.2.5).



Figure 3.15 Index-finger pointing overlapping with lip-pointing.

immediately follows this (in Figure 3.14c) with an index-finger point to the same location (with neither aligned gaze nor accompanying speech).

His interlocutor is a close relative who once lived in the same house (i.e. the house where the interview is taking place), but who has not been there for some time. The speaker's initial lip-point, with its element of 'I think you know where I mean', is 'up-graded' here with an index-finger point, due to the speaker's realizing that his interlocutor, no longer living in the house, probably indeed does **not** know where he means.

The next example (Figure 3.15) is of similar structure, although the lip- and index-finger pointing are more closely aligned, the lip-point only fractionally preceding the index-finger point, and the two pointing gestures being mostly simultaneous. The speaker has just returned from a fruitless fishing expedition, and is discussing the relative merits of different local fishing spots. He is asked where he has just been fishing. He responds (in Figure 3.15a) by immediately beginning a lip-pointing gesture in the direction of the referent location (out of sight, about 200 metres away), turning his head towards the referent, and beginning to utter the demonstrative *nuun4*, a variant of the far distal adverbial *phuun4* 'yonder'. At this moment, his hands are occupied, and it is conceivably because of this that he is shaping to lip-point. But given that the local fishing spots are only a limited set, it may be quite natural for him to use a lip-point anyway, to signal an assumed familiarity ('I think you know where I mean'). However, his interlocutor is not a local, and certainly could not be assumed to be familiar with the various fishing spots around the village. It appears that once the lip-pointing action has begun, the speaker realizes that a lip-point alone could be presumptuous (since his interlocutor probably **doesn't** know where he means), and he decides that an index-finger point is required. Note that he does not abandon his lip-point already underway. Within three video frames (around 100 milliseconds) of starting the lip-point action, the index-finger point begins, and is accompanied



Figure 3.16 (a) ‘Look, (it’s) here!’; (b) ‘The big tall house, here!’

throughout by the lip-point, until reaching full arm extension (Figure 3.15b), when both pointing gestures are retracted together, head returning to initial position.

In a few cases, lip-pointing is immediately **preceded** by index-finger pointing. In the next example, the speaker has been asked where a certain person lives. The speaker first index-finger points (Figure 3.16a) in the direction of the person’s house (out of sight, approximately 300 metres away), with fully extended arm, held for nearly one second, as she says *beng1 ndêê4* [look PCL] ‘Look, (it’s) here!’ Then, having retracted her index finger-point (in Figure 3.16b), she further comments *hùan2 ñajø-ñaj1 sungø-suong3 ndêê4* [house RDP.A-big RDP.A-high PCL] ‘The big tall house, here!’, lip-pointing as she speaks.

The same kind of explanation as offered for the opposite sequence (lip-point followed by index-finger point), above, may be offered in this case, but in reverse. The speaker first responded by index-finger pointing, apparently the default form of spatial deictic gesture in the Lao system. However, in this case the speaker knows that the interviewer is very familiar with the village, and is certainly familiar with the house in question, which is large and in a prominent location. I suggest that because of this, the speaker re-phrases with a lip-point, to indicate that her previous neutral response (with index-finger point) was inappropriate. The recognitional meaning of the lip-point (‘you know which one I mean’) is more appropriate in this case. Linguistic evidence also supports this analysis. The speaker’s linguistic encoding of the referent at the time of the lip-point involves a particular type of reduplication, also of a ‘recognitional’ nature (specifying that the referent is assumed to be discernable by the listener among a range of possibilities; see Enfield 2007: 253–255).

These observations about the function of lip-pointing in the context of its sequencing with other deictic gestures demonstrate that different pointing

gestures are not merely interchangeable alternatives. Different deictic gestures have distinct functions, and these functions are exploited in complex ways in the enchronic unfolding of discourse.

3.2.3 *The gaze switch hypothesis*

As far as can be determined in the Lao data, lip-pointing is always accompanied by aligned eye gaze in the direction of the referent. Discussions of examples illustrated in Figures 3.1–3.5, 3.9–3.11, 3.14–3.16, above, provide details of the pattern of alignment of eye gaze with the lip-pointing gesture. These observations give rise to a gaze-switch hypothesis. It is well-documented that people are highly attuned to the direction of gaze of their social associates (Gibson and Pick 1963, Cline 1967, Baron-Cohen 1995, among many others), a fact of special importance in joint attention, and human interaction in general (Tomasello 1999, Enfield and Levinson 2006). Accordingly, I suggest that the primary vector-projector in lip-pointing gestures is not the lips but the eye gaze, and that the role of the lips in these gestures is to index the deictic function of gaze, to trigger or ‘switch on’ the deictic vector that eye gaze projects. Gaze can point but cannot point all the time. For an interpreter to understand that an interlocutor is pointing with his gaze, there needs to be a trigger that he is at that time intending to indicate something with his gaze.

Thus, the lip-point is a handy way of triggering the deictic function of gaze, a signal saying ‘I’m now indicating something to you by looking at it.’ The same mechanism applies for other head-area spatial deictic gestures such as head-toss, chin-jut, eyebrow-raise, and bug-eye gestures which all typically accompany lip-pointing. In support of this hypothesis are examples of lip-pointing in which gaze and face/lip orientation are not precisely aligned (cf. Figures 3.2 and 3.4, above). In these cases it is clearly the direction of the gaze, not of the face (nor the lips, since lips can hardly orient at an angle away from the face), which we take to be pointing to a referent.

This is in distinct contrast to the use of hand-pointing, in which the speaker’s gaze need not be aligned with the direction of the deictic gesture at all, being instead employed for other higher-priority purposes, such as management of interaction between interlocutors. Note Figure 3.14c, above, for example, in which the speaker is finger-pointing while his gaze is engaged with his interlocutor, while in the lip-point immediately preceding (Figure 3.14b), gaze is aligned with the lip-point. Let us consider an example which illustrates that hand-pointing can allow the option of not aligning one’s gaze with one’s pointing gesture, when gaze is required for other purposes. Figure 3.17, from some two minutes earlier in the same sequence as Figure 3.13, above, shows the speaker (second from left) pointing



Figure 3.17 Mismatch of direction of gaze and index-finger pointing (cf. Figure 3.13, above).

in the direction of an out-of-sight referent in the village (a primary school some 500 metres away), and his extended index-finger point is not accompanied by gaze in the direction of the referent. Rather, his gaze is directed towards his interlocutor.

In Figure 3.13, above, the same speaker makes the same index-finger point to the same referent, yet on that occasion aligns his gaze with the gesture. The difference between the two cases concerns the relationship between the speaker, his addressee (to the speaker's right), and the current conversational state of play. In Figure 3.13, the speaker was one of the two comprising the interactional dyad, which is currently dominating the floor. He was already engaged in conversation with the investigator, to his right (the other two men were listening and not interjecting), and did not need to use gaze for any handling of interaction at this point. By contrast, in Figure 3.17, the speaker is **not** already holding the floor. In this case, a dialogue is taking place between the investigator and the man to the speaker's left (wearing a white shirt). The white shirt man has just answered the question 'Where is the village school?' posed by the investigator, and in Figure 3.17 the speaker (second from left) interjects with a parenthetical clarification, adding some elaborating information to what the man on his left has just said. This speaker is not claiming the floor, merely inserting a remark for the benefit of the investigator, and for this reason he keeps his gaze focused on his addressee as he carries out the pointing gesture (i.e. signalling 'this is just between me and you'). In this

case, the speaker's gaze is available to be used independently of the deictic hand gesture for higher priorities in the interaction. While this kind of non-alignment between gaze and the direction of the deictic gesture is common with hand-pointing, it is unattested in the lip-pointing examples.

3.2.4 *Place of lip-pointing in the deictic gesture system*

In the context of the observation that index-finger pointing 'is sometimes suppressed in some cultures' (cf. Wundt 1973 [1921]: 75), Hewes (1981: 265) refers to lip-pointing as 'an occasional alternative to finger-pointing'. But lip-pointing should not be negatively defined, i.e. as a mere substitute for a suppressed finger-point. It has been shown above that lip-pointing and index-finger pointing can occur in the same contexts (see Figures 3.14–3.16). Lip-pointing holds its own place in a larger system, and is not to be defined merely in terms of a residual contrast with hand-pointing, especially given that there are so many forms and functions of the latter. Lip-pointing needs to be treated on its own terms, for at least three reasons.

First, the data from Lao show that lip-pointing is used when there is no reason to imagine that index-finger pointing is suppressed (*pace* suggestions to the contrary; e.g. Hewes 1981). Lao speakers often claim that lip-pointing is what you do when your hands are not free for hand-pointing. Similarly, the lip-pointing gesture among Navajo speakers is said to be 'quite handy when someone asks you a question and your hands are full' (DiLucchio 1998/1999). True, but this is more likely a convenient fact than a genuine motivation. People consistently lip-point when their hands are free (see examples in Figures 3.2, 3.3, 3.11b, 3.12b, and 3.16b, above), and furthermore they use both lip- and hand-pointing during the same communicative scenario, sometimes simultaneously or together in the same utterance (see above, especially Figure 3.15b).

Second, just as hand-pointing (or index-finger pointing more specifically) may be culturally suppressed in certain circumstances, so also may lip-pointing. Given that, as Sherzer (1993: 195) points out, lip-pointing can be associated with lower social standing, then it is just as likely that index-finger pointing is motivated by a suppression of lip-pointing. It is widely assumed that index-finger pointing is 'basic' (cf. discussion in Wilkins 2003), but this remains an empirical question, as we await greater breadth and detail of cross-cultural data on deictic gesture systems. So, while lip- and index-finger pointing may in some sense be alternatives, they are not to be **defined** as such. They each have their own value in the system.

Third, lip-pointing may carry specified meanings beyond a basic deictic function, and its semantics will thus require positive characterization, not mere allocation to a residual category. Sherzer's (1973) data from the Kuna people of Panama pays special attention to meanings beyond basic spatial

deixis (including interpersonal joking and mocking functions), leading us also to a variationist perspective on gesture, to which we now turn.

3.2.5 *Sociokinesics and gestural register*

Lip-pointing by Lao speakers is low register or sub-standard. When discussing lip-pointing with Lao speakers, one elicits the kinds of response one would expect in discussions of bad language or rural dialect linguistic forms, namely, amusement, embarrassment, and quizzicality as to why a scholar would be interested in the phenomenon at all. In correspondence about reported lip-pointing among Navajo speakers (for whom lip-pointing is customarily defined as a preferred option to index-finger pointing, since the latter is considered impolite), Larry DiLucchio of the Navajo Nation wrote:

I shared your letter with my office staff. Most were very amused that you would take such an interest in something so mundane . . . A photograph? Perhaps, if I can get them to quit laughing at the thought.

One explanation of this low register status of lip-pointing is the following. Both lip-pointing and index-finger pointing can serve a basic spatially deictic function, i.e. of simply indicating the location or existence of something in the immediate environment. If both forms of pointing become conventional in a culture, then they will not contrast in this function, and thus the basic deictic function of one of these is likely to become enriched in some way as to produce a contrast, and thereby both justify and maintain the existence of both forms in the one system. One logical possibility is for the two to continue to perform the same referential function (basic spatial deixis), but split in terms of social register. This would account for the social values of lip-pointing found in Laos and elsewhere. It is interesting to consider why it is that when lip- and finger-pointing are socially distinguished, lip-pointing tends to be regarded as more crude or uneducated. Perhaps there is something more animal-like about the unmediated instrumental use of the face (cf. Vygotsky 1962 [1934]).

A second possibility is for one of the forms to take on an enriched semantic value. For example, the recognitional uses described here for Lao, the mocking, ironic meanings of lip-pointing among the Kuna, described by Sherzer (1973), or more fine-grained semiotic distinctions, as different forms of hand-pointing may have (Kita 2003, Kendon and Versante 2003, Wilkins 2003).

3.2.6 *Deictic gesture system typology*

Among Lao speakers, lip-pointing and many forms of pointing by hand coexist. No form of spatial deictic gesture seems secondary or routinely

suppressed in favour of others, and the system as a whole appears to be highly structured. Wilkins (2003) describes a similar degree of structured systematicity among forms of deictic gesture, including lip-pointing and a range of hand-pointing forms, by Arrernte speakers of Central Australia (cf. also Kendon and Versante 2003 for description of systematic variation of hand-pointing forms by Neapolitans; see Chapter 4, below, for systematic variation by Lao speakers). An important typological question arises: Given that there are deictic gesture systems with conventions of hand-pointing but no convention of lip-pointing (such as is the case among English speakers), are there any with conventions of lip-pointing but no conventions of hand-pointing? In some cases it seems that finger-pointing is less common or less salient than lip-pointing. Sherzer, for example, says that ‘hand and finger pointing occurs infrequently’ among Kuna speakers, and that lip-pointing is ‘more common by far’ (Sherzer 1983: 169). Wilkins (2003: 207) suggests that there are gesture systems in which ‘the canonical form of pointing is a lip point and there is no, or little, evidence of systematic pointing with the index finger’. This intriguing possibility needs empirical support. Wilkins’ statement is restricted to index-finger pointing rather than hand-pointing more generally, and it seems unlikely that a culture could be found without some form of systematic and conventional pointing by hand. We need more information about such systems in cultures around the world.

3.3 Concluding remark

Lip-pointing among Lao speakers shows consistent conventions of both form and function. Lao speakers mostly lip-point with lips parted, often with the upper lip primarily protruding. Also, in most if not all cases, a lip-pointing gesture involves other kinesic actions, especially tilting the head up, sticking the chin out, and opening up the eyes (with eyebrow-raise). Patterns of coordination between lip-pointing and other forms of deictic gesture (especially hand-pointing of various kinds) can reveal functional distinctions. Lip-pointing by Lao speakers is distributionally restricted to utterances in which the location or identity of the referent is in focus in the utterance. These gestures almost exclusively appear as answers to ‘Where?’ or ‘Which one?’ questions. Index-finger pointing, on the other hand, shows much wider distribution (cf. Chapter 4, below). Another functional distinction of lip-pointing is its recognitional sense (‘I think you know where/what I mean’), which accounts for its informal/familiar – and thus low register – status in the Lao system. Finally, the observation that lip-pointing among Lao speakers is always accompanied by gaze directed toward the referent suggests that the deictic vector is not provided by the action of the lips, but by the gaze. The action of the lips indexes that the speaker is pointing with his or her gaze at

that moment. To be a form of pointing, gaze needs to be overtly triggered, since we cannot assume that our interlocutors are pointing something out whenever their eyes are open.

While this chapter has concentrated on data from Laos, it is hoped that the observations may stimulate further cross-cultural comparison. Basic descriptive work on the form and function of lip-pointing around the world is scarce. For any lip-pointing system, we need information about: (a) the conventional form(s) of the lips in such gestures; (b) the possibilities and likelihood of accompanying kinesic actions, involving head, eyebrows, chin, and gaze, as well as speech, including the relative timing and functional interaction of these; (c) the precise range of deictic and other semiotic functions which lip-pointing may perform, with special attention to the relationship between the deictic act and the discourse/communicative context; and (d) the place of lip-pointing in the larger structured system of conventional gesture, in terms of both its semantic and social functions.

4 Hand-pointing¹

If one has the attitude of the person throwing the ball, he can also have the response of catching the ball.

G.H. Mead, 1934

Despite long being considered the simplest, most primitive form of communicative action (Wundt 1973 [1921]: 74–75), the humble pointing gesture has much to reveal about the nature of human social intelligence and mind. For nearly a century, scholarly considerations of pointing gestures have addressed questions of the very foundations of signification, from psychologists' ruminations on their production and comprehension (Wundt 1973 [1921]: 74–75, Bühler 1982 [1934]: 45), to philosophers' concerns with notorious puzzles of ostensive reference (Wittgenstein 1953: §33, Quine 1960: 100–101; cf. Peirce 1965 [1932]: §2.305), to pioneering ethnographic descriptions of pointing's role in different cultural settings (Efron 1972 [1941]: 96–97, Sherzer 1973, Haviland 1993). In more recent decades, most work on pointing has focused on its cognitive and semantic status as a proto-linguistic mode of human communication. Pointing is argued to be a critical precursor to language, both in phylogeny (e.g. Hewes 1981, 1996, Rolfe 1996) and ontogeny (Scaife and Bruner 1975, Bates *et al.* 1975, Ohama 1984, Schaffer 1984, Bates *et al.* 1987, Carpenter *et al.* 1988, Butterworth and Morissette 1996, Butterworth 2003, *inter alia*). Pointing is now understood to constitute a uniquely human mode of joint-attentional behaviour (Povinelli *et al.* 2003, Tomasello 2006). Recent experimental work with prelinguistic infants implicates pointing directly in one of the defining achievements of the human mind, our capacity for shared intentionality (Liszkowski 2006, Tomasello *et al.* 2005). But the interest in pointing here goes beyond its

¹ The study described in this chapter was carried out together with Sotaro Kita and J. P. de Ruiter (see Enfield, Kita, and de Ruiter 2007, *Journal of Pragmatics* 39, 1722–1741). Thanks to Kita and J.P. for agreeing to the revision and inclusion of the chapter here. The original impetus came from field research materials designed by Sotaro Kita and David Wilkins in 1999 in the Language and Cognition Group at the Max Planck Institute for Psycholinguistics, Nijmegen. Updated versions of these materials can be obtained from the author.

proto-linguistic status. After prelinguistic infancy, pointing goes on to play a significant role in fully linguistic interaction. That is, once children acquire language, they do not stop using the pointing gesture as a key tool for achieving joint attention and constructing composite utterances.

Pointing with the hand is a basic mode of sharing attention and constructing composite signs that incorporate token environmental structure. The basic-level status of pointing is modulated in adult behaviour by the fact that different practices emerge within conventional systems, and show meaningful contrasts, analogous to the systematic contrasts in linguistic deixis discussed in Chapter 2. This chapter presents an analysis of two types of spatially-anchored pointing gesture which form composite utterances with accompanying speech in two very different ways. In one type – B-points, ‘big’ in form – the information in the hand movement is a primary, foregrounded component of the utterance, along with accompanying speech and sometimes even without it (e.g. when I silently point to your keys in response to your question ‘Where are they?’). B-points are equivalent to proto-linguistic gestures, and, unlike many types of hand gesture, are not necessarily dependent on language for their proper interpretation. A second type – S-points, ‘small’ in form – shows a different division of labour: the speech is informationally foregrounded, the hand movement is a backgrounded modifier. These smaller, backgrounded pointing gestures are of an entirely different degree of informational sophistication from B-points. S-points are functionally stitched in to spoken utterances. What look at first to be slight variations in form (bigger versus smaller pointing gestures) turn out to indicate qualitatively quite distinct types of phenomenon, with different underlying cognitive and communicative structures. Evidence suggests that they are acquired at opposite ends (i.e. first, and last) of the order of acquisition of gesture types. That is, what look at first sight like *a* and *b* turn out to be alpha and omega.

In what follows, I first describe an empirical study of Lao speakers’ spatially-anchored pointing gestures, accounting in functional terms for the differential distribution of B-points and S-points (formally defined below). The question is whether the variation between big versus small form of a spatially-anchored pointing gesture is random, or whether this formal variation is related to the pragmatic function of the composite utterance in which it occurs. The data suggest that form/function differences between the two types of pointing gesture reflect distinct types of constraint which interactants have to satisfy in confronting the online problem-solving tasks of designing moves in face-to-face interaction. The form/function properties of B-points are relatively straightforward, seemingly the same as the infants’ earliest communicative gestures (Liszkowski 2006). They focally serve to establish joint attention, and to inform, with respect to some perceptually available, and spatially locatable, object. By contrast, S-points have more

subtle information-structural properties. These gestures arise where an informational imperative comes into direct competition with a social-affiliational imperative (Enfield 2006). These two imperatives are among multiple elements contributing to a speaker's constraint satisfaction problem in the course of producing composite utterances for communication. The informational imperative requires interactants to ensure that the calibration of their referential understanding is sufficient for current communicative purposes at any stage of the interaction – or, in slogan form, 'Don't under-tell' (Schegloff 2007a). The requirement for such calibration is to be expected in an intelligent system of communication which follows an enchronically cumulative informational trajectory, and in which signals are designed for their intended audiences (Mead 1934, Sacks and Schegloff 1979/2007, Clark 1992, 1996). But at the same time, since every exchange in human interaction is socially consequential (Sacks 1992, Heritage and Raymond 2005, Enfield 2006), we can expect utterance production to be as much constrained by social-affiliational imperatives as it is by informational imperatives. We learn from S-points in the Lao data that social-affiliational imperatives outrank informational ones when the two are in conflict. In later sections of the chapter, we consider related phenomena described in other cultural settings, and we look more broadly at these secondary, embedded functions of pointing gestures within linguistically complex composite utterances.

4.1 A controlled study of hand-pointing

This chapter describes a controlled study of a set of hand-pointing gestures during speech by Lao speakers, taken from a corpus of conversation and semi-structured interviews in villages in and around Vientiane.² From these recordings, we isolated all composite utterances which featured a spatially-anchored pointing gesture. A pointing gesture was defined as a communicative bodily movement which projects a vector whose direction is determined, in the context, by the conceived spatial location, relative to the person performing the gesture, of a place or thing relevant to the current utterance. Excluded from consideration were those pointing gestures which featured (a) hand shapes other than those with a single protruding digit (since it is known that these can have special, elaborated communicative functions such as politeness, denoting features of the target etc.; Kendon and Versante 2003, Wilkins 2003), or (b), arc or tracing movements of the articulator at phase of full extension, apparently representing motion (e.g. as occurred in our

² Included among these materials are video-recordings of the Historically-Anchored Locality Description Task, a semi-structured outdoor interview on the history of a locality (Wilkins *et al.* 2001).

data in references to local roads), since other factors are likely to be involved in their use. All 102 gestures isolated by this procedure featured origo and target on the same horizontal plane. That is, none of the gestures pointed to a referent located above or below the speaker.³ They all involved articulation of the gesture by one arm only (no two-handed deictic gestures occurred), and all featured a single digit (index finger or thumb) as vector-projecting articulator.

Each utterance in the set was coded for form of the gestural action.⁴ Coding was for manner of articulation of the gesture, distinguishing formally between B-points (gestures in which the whole arm is used as articulator, outstretched, with elbow fully raised), and S-points (gestures in which the hand is the main articulator, the arm is not fully straightened, typically with faster and more casual articulation). Also coded was orientation of the head, towards versus not towards the target. Information about head orientation allowed for inference of direction of eye gaze, a deictic signal of obvious importance in pointing (see following section for discussion). The two types of gesture occurred with similar frequency in our set of cases (forty-six B-points vs. fifty-six S-points). Examples are shown in Figures 4.1 and 4.2.

The question is what the functional distinction is, if any, between these forms of co-speech pointing in the Lao corpus. It was found, first, that B-points are characteristic of location-focus utterances (see section 4.1.1, and Figure 4.8, below, for discussion and evidence). Let us first discuss these cases, before turning to S-points.

4.1.1 *B-points in location-focus utterances*

The term location-focus describes a property of composite utterances in which the location of the referent is focal in the sense of Lambrecht (1994: ch. 5), and nothing in the linguistic component of the utterance imports a presupposition of specific location of the referent (cf. Chapter 3, above). In other words, the primary function of the utterance is to tell the addressee where something is, in such a way that the ‘where’ information is given by token, non-conventional signs in the situational context (possibly including a gesture), not by the speech (see Figure 4.1). This typically means that a symbolic indexical element in the speech, such as *here* or *there*, obliges a listener to go looking in the context for what is intended. Often it is a hand

³ Also, none of the gestures pointed to a referent within physical reach. Had such gestures been observed in the data set, they would have been excluded, since it is known that the form of a pointing gesture can be formally adapted in special ways to the fact that a referent is small or nearby (independent of information structure considerations).

⁴ Coders did not understand Lao, and were not aware of the research question.



Figure 4.1 B-point (whole-arm articulation) with head oriented to target (speaker – in white shirt – pointing to something forward and to his left).



Figure 4.2 S-point (hand/lower-arm only articulation) with head not oriented to target (speaker – in white shirt – pointing to something directly behind him, protruding thumb obscured).

gesture that directs the addressee to the key information, but it need not be. There are other means to fill in the blank that a deictic term opens up, the key mechanism being that a solution is made mutually salient (Clark *et al.* 1983).⁵

In all of the location-focus utterances in the data set which featured a pointing gesture (thirty-four in total), that gesture was a B-point (see Figure 4.8). The canonical case is when an utterance answers a ‘where?’ question with a demonstrative like ‘here’ or ‘there’ in combination with the B-point gesture:

- (1) I *juu1 thaang2 daj3 hoong2-hian2*

be.at way INDEF school

‘Where is it, the school?’

D *hoong2-hian2 juu1 thaang2 [suan3-maak5 thaang2 phi4*
school be.at way garden-betel way right.here
dêø]

FAC

‘The school’s in the direction of [the betel garden this way *right here*]_{Fig. 4.1}’

The gesture by D in (1), coinciding with the square-bracketed section of speech, is illustrated in Figure 4.1.

In another kind of example of a location-focus utterance, the locational information is volunteered independently by the speaker (rather than supplied in response to a ‘where?’ question). The utterance similarly has the primary function of telling the addressee where something is. In this example, the speaker is talking about the history of his locality, and is discussing various *tasèng3* ‘district chiefs’ who had formerly governed:

- (2) *tasèng3 ngoon1 [khaaw1 nan4, naa2 phen1 ñang2 juu1*
district.chief N period that, paddy 3SG.P still be.at
nii4 dêø –]

here FAC.FILLIN

‘The District (Chief) Ngoon [at that time, his paddy fields were still *here*]_{Fig. 4.3}’

⁵ In defining location-focus utterances for current purposes, this excludes cases in which the speech itself imports a presupposition of location of the referent, the main type of case being the use of proper names. Thus, *It’s in Sunnydale* relies on the addressee already knowing where Sunnydale is, while *It’s there* relies on the addressee to recognize in the context a mutually salient solution to the problem of where ‘there’ must be referring to. Utterances such as *It’s in Sunnydale* do focus on location, but these do not concern us here. Location-focus utterances which include symbolic indexicals require reference to context, but note that they do not entail the use of a hand gesture. Nor do they specify, if a gesture is used, what form it should take. This question is considered in the following section.



Figure 4.3 ‘The District (Chief) Ngon [at that time, his paddy fields were still here].’

The use in speech of a demonstrative such as *nii4* ‘here’ or *phii4* ‘right here’ requires the addressee to consult context in order to determine the intended referent. A common aid to this search is an accompanying pointing gesture. De Ruiter and Wilkins (1998: 605) call these ‘obligatory’ gestures, since the spoken part of the utterance ‘requires a gesture to go along with it for it to be codable and semantically complete’.⁶ These are distinguished from ‘supportive’ gestures, in which the spoken part of the utterance ‘could stand on its own in context (with the primary proposition intact)’ (ibid.; see section 4.1.2). The ‘obligatory’ gesture helps determine the mutual salience required for a speaker and addressee to converge on the intended referent of a symbolic indexical such as a demonstrative (e.g. English *this*; Clark *et al.* 1983; cf. Chapter 2, above). This chapter is concerned with the relation between two forms of pointing gesture (B- and S-points) and their contrastive semiotic function. While the notions of obligatory and supportive are a useful start, we shall try here to focus on the qualitative difference between informational roles that gestures can play in composite utterances, whether formal differences align with these, and if so, why.

⁶ To say that the gesture is literally obligatory is too strong, since the reference of a deictic can be found in the context through other means, as long as a mutually salient solution is available (Clark *et al.* 1983).

When a referent is spatially distant, both B-points and S-points occur. But for an utterance with a location-focus function it is a B-point that serves, articulated with the whole arm, and usually with head and gaze also oriented to the target. This ensemble of features is characteristic of infants' primordial pointing gestures whose function is to inform or otherwise purposely establish focal joint attention (Lizskowski 2006, Tomasello 2006). Let us consider why it is that these location-focus composite utterances feature B-points.

4.1.1.1 Why B-points are formally well fitted to location-focus contexts Location-focus utterances have the primary purpose of telling an addressee where something is. B-points serve this function better than S-points for at least three reasons. First, they are better for attention-direction, because they take up more of the addressee's visual field. They move further, and they move faster. Second, the greater physical effort involved in a B-point is a potential index of greater communicative importance of the locational information (i.e. the information expressly conveyed by the gesture) to the primary assertion of the speaker's utterance (Levy and Fowler 2000). Third, B-points are maximally accurate as pointers (in the absence of long sticks or laser pointers). They have longer, straighter vector-projecting form, and are typically held for longer duration. These three properties of B-points make them better fitted for conveying focal locational information. In location-focus utterances, where context is critical for comprehension of locational reference, the B-point does it best.

In addition to these location-focal features of the form of B-points' hand/arm articulation, B-points also tend to occur with accompanying orientation of head to target, significantly more often than S-points do (Fisher's exact test, $p < 0.001$). In addition to coding of gesture 'size', we coded for orientation of the head either to, or away from, the target of the pointing gesture at the peak of the gesture's articulation. The two binary parameters of manner of S- vs. B-point articulation and towards vs. away head-orientation yield four formal variants, distributed in the data as shown in Figure 4.4.

Two generalizations are possible:

- (3)(a) When a B-point is performed, the head will likely be oriented to the target of the pointing gesture ($39/46 = 85$ per cent). (By contrast, S-points only weakly prefer head-not-to-target.)
- (b) When the head is not oriented to target, a pointing gesture will likely be an S-point ($32/39 = 82$ per cent). (By contrast, head orientation to target only weakly prefers B-points.)

In short: if the pointing gesture is a B-point, then head (and gaze) will be oriented to the target of the gesture. It is well established that the direction of eye gaze (inferable from head orientation in the absence of reliable eye-gaze

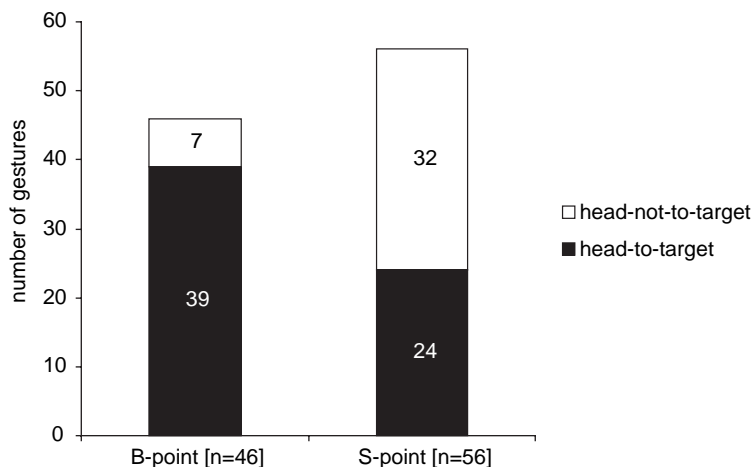


Figure 4.4 Distribution of B-point (whole arm) and S-point (hand/lower-arm only) gestures to orientation of head to target.

data) is a powerful component in attention direction (Baron-Cohen 1995; cf. Chapter 3, above).

To summarize, the formal features of B-points, along with the aligned eye gaze which typically accompanies them, account for their good fit with the function of focusing on ‘where’, and therefore on drawing focal attention to ‘where’ at the utterance level. The kinds of utterances which include these gestures are informationally straightforward. They are not much different to the proto-linguistic informing utterances that twelve-month olds perform with pointing gestures (Liszkowski 2006). (This is of course not to deny the difficulties of interpreting ostensive gestures, regardless of whether they are accompanied by speech; Wittgenstein 1953: §33, Quine 1960: 100–101, Goodwin 2006.) We now turn to the more complex matter of S-points.

4.1.2 *S-points in insecure-reference environments*

In contrast to the straightforward distribution of B-points to a single utterance function, the kinds of situations in which S-points occur are more varied. Here are the four classes of utterance type which feature S-points in the data set:

- (4)(a) Request for clarification: the speaker is asking for clarification that this referent was the intended referent of a previous utterance (e.g. *You mean [the school]_{S-POINT} for secondary students?*)

- (b) Supply of clarification: the speaker is clarifying/confirming that this was the intended referent of a previous utterance of their own; in response to addressee's request for clarification (e.g. *What I mean is [the school]_{S-POINT} for secondary students.*)
- (c) Very general reference form: the speaker's linguistic utterance appears to contain insufficient information for unambiguous recognition of the intended referent, e.g. due to ellipsis or pronominalization (e.g. *[It's]_{S-POINT} for secondary students.*)
- (d) Initial mention of a recognitional term (e.g. a place name): the speaker introduces the referent with a proper name for the first time in the discourse, where there is possible reason to think that the linguistic term alone might not be sufficient for reference to be secured (*My kids go to [Sunnydale School]_{S-POINT}.*)

There is a functional unity to these seemingly diverse utterance types. In each case, the speaker makes linguistic reference to a uniquely recognizable entity in a situation where the referent seems likely but not certain to be recognizable for these interlocutors in this context. There is apparent (if low) risk that the addressee is not going to be able to achieve reference on the basis of what's being said in the talk. These may be referred to as insecure-reference environments. The best way to clarify is with examples.

4.1.2.1 Three examples of S-points in insecure-reference

environments S-point with request for clarification: 'The steel bridge'.

In this example, the interviewer wants to ask about a particular bridge he has in mind ((5), line 1), not aware that there are multiple bridges he could be taken by this recipient to be referring to. The utterance in line 1 is insufficient to secure common reference to a single bridge in the local area. This is made apparent in line 2, when a stretch of more than a second's silence reveals an absence of uptake or grounding of the interviewer's attempted reference (Jefferson 1989, Clark 1996: 268). In line 3, the interviewer resumes, voicing a signal of disfluency. Here, the addressee takes over and begins a request for clarification (lines 4–6), proposing a candidate referent for the bridge referred to by the interviewer in line 1. His S-point (illustrated in Figure 4.2, above) co-occurs with his spoken description of this candidate referent in line 6, the 'steel bridge':

(5)

- 1 I: *suan1 qanø khua3 juu1 thaang2 lum1 hanø dêø*
 as.for HES bridge be.at direction below there FAC.FILLIN
 'As for, um, the bridge down below there?' [No gesture]

- 2 (1.3s)
 3 I: *qanø*. . .
 ‘Um . . .’
 4 D: *khua3*. . .
 ‘(The) bridge . . .’
 5 (.3s)
 6 D: *khua3* [*mak-lêk1* (.3s)] *khua3 phaani1 maa2 niø vaa3*
 bridge steel bridge cross come TPC QPLR.INFER
 ‘The bridge of [steel]_{Fig. 4.2}? The bridge (you) pass coming (here),
 you mean?’

The S-point marked in line 6 with square brackets is timed to peak with articulation in speech of the identifying feature of the bridge – i.e. that it’s the steel one, and not one of a number of other bridges made of other material, in other locations near the village.

S-point with very general reference form: ‘Ø not find (it) delicious?’

Two women are sitting at an outdoor table (Figure 4.5; Da is on the left, Tui on the right). They are eating a snack food which has been shared around to them along with a number of children and visitors who are milling about. The discussion turns to whether the children are enjoying the food. At a certain moment, Da formulates a composite utterance whose ambiguity is foreclosed only by the S-point (directed at her addressee Tui)⁷ which accompanies it (Figure 4.5):

- (6) *bòø* [*sèèp4*] *vaa3*
 NEG find.delicious QPLR.INFER
 ‘(Ø=You) don’t find (it) delicious?’

The spoken part of this composite utterance is ambiguous because the subject of *sèèp4* (here meaning ‘to find something delicious’) could be the addressee Tui, or a child Khim, who is currently topicalized in the conversation. The pointing gesture disambiguates the overly general reference form in the utterance it accompanies, by filling in the gap in linguistic structure left by the zero anaphor (cf. Levinson 1987: 114–118; Haviland 1993: 17). Examples like this make explicit one of the key cognitive challenges in both producing and comprehending composite utterances: namely, binding conventional and non-conventional signs (here, speech and gesture) into a unitary informational structure at the level of the communicative move.

⁷ Since the gesture is performed in the casual, fleeting manner typical of S-points, it is difficult to depict the gesture in a still image captured from video. Figure 4.5 shows Da’s hand at its fullest extension, but the hand is in motion at the moment depicted.



Figure 4.5 ‘(Ø=you) don’t find (it) delicious?’

S-point with initial mention of recognitional: ‘Saythani District’.

The interviewee is describing changes that have taken place in his locality over the years:

- (7) *tèè1 kòòn1 mùang2 baan4 haw2 niø*
 from before town village I.FA TPC
ñang2 khùn5 kap2 mùang2 [saj2thani2]
 still depend with district Saythani

(pause)

bòø dajø khùn5 kap2 mùang2 [saj2sêêt5thaa3] dêj2
 NEG ACHV depend with District Saysettha FAC.NEWS

‘Before, our town and village still depended on [Saythani]_{Fig. 4.6} District. It didn’t depend on [Saysettha]_{Fig. 4.7} District, you know.’

The first pointing gesture (Figure 4.6) is to a place behind the speaker, to his left. The second gesture (Figure 4.7) is to his right, forward. Both gestures are S-points. The spatial locations to which these gestures aim are where the offices of the district administrations are physically located (each several kilometres away). These references are insecure because while the speaker has reason to believe that his addressee will know what is meant by the place names *Saythani* and *Saysettha* (the interviewer is known by the speaker to have been familiar with the area for some years), there is some reasonable doubt that these terms will be fully understood (since the interviewer is, after all, not a local).

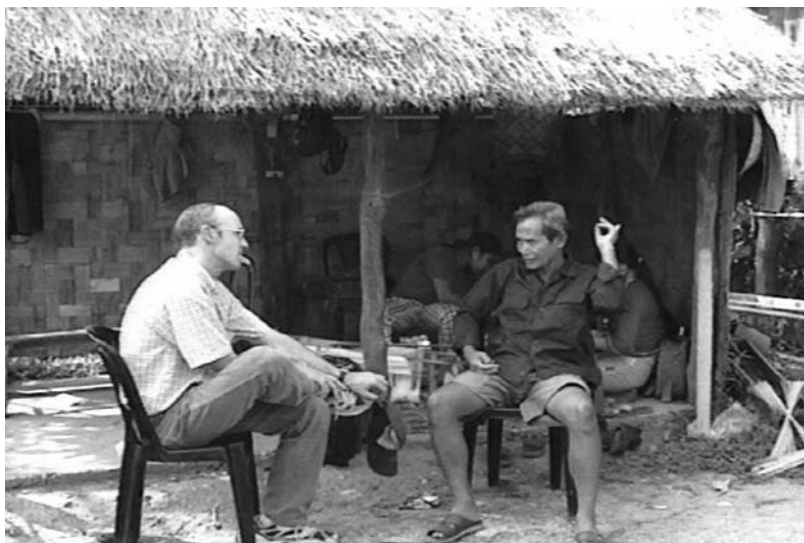


Figure 4.6 'Before, our town and village still depended on [Saythani] District.'



Figure 4.7 'It didn't depend on [Saysettha] District, you know.'

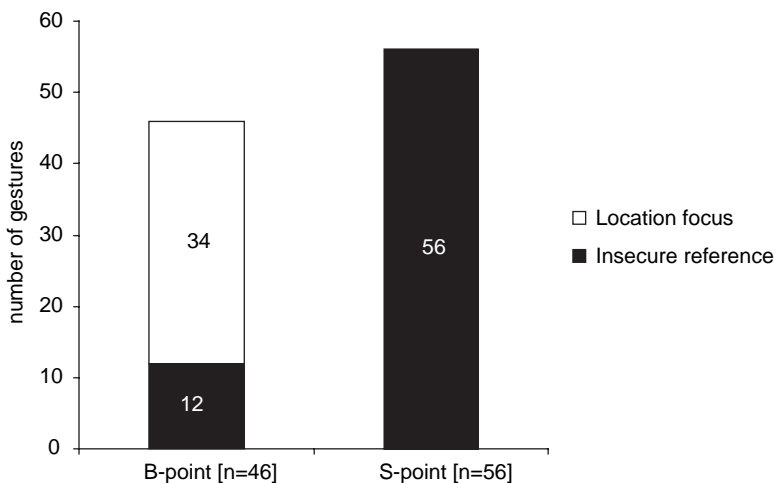


Figure 4.8 Distribution of B-point (whole-arm) versus S-point (hand/lower-arm only) in the two utterance functions. (No S-points occurred in the location-focus condition.)

4.1.2.2 When informational and social imperatives conflict: the S-point as a safety net for over-supposition Of the set of 102 pointing gestures, around two-thirds (sixty-eight) occurred in insecure-reference contexts. Of these, more than four-fifths (fifty-six) were S-points (a significantly greater proportion than in the location-focus context, where S-points were not observed; Fisher's Exact test, $p < 0.001$). Further, B-points occurred more often in location-focus utterances than in insecure-reference utterances ($\chi^2(1) = 5.26$, $p < 0.05$). In other words, S-points are associated with insecure-reference utterances, and B-points with location-focus utterances.

These S-points are selected in a class of interactional situations in which speakers face a dilemma: 'I want to formulate a likely identifiable referent as recognitional (e.g. *You come in past Somsamai Village*), but I'm not completely confident that you will recognize the referent by the linguistic formulation I have in mind.' The speaker is stranded between options for linguistic formulation. If her linguistic formulation supposes too much – e.g. by introducing a place by proper name (as in example (7), above) when it turns out the recipient doesn't recognize the place by name – the risk is failure to achieve reference. This fails to satisfy an informational imperative, resulting in failure of speaker and addressee to converge in referential understanding. Under-telling in this way is likely to result in a request for (and subsequent supply of) clarification, disrupting the progressivity of the

interaction (Stivers and Robinson 2006), at some cost for all concerned (cf. Heritage 2007, Schegloff 2007a).

Suppose that the only operative imperative in an intelligent system for communication were to provide information sufficient for referential success. In that case, if a given referential formulation were thought by its producer to be possibly inadequate for its intended recipient, it should simply be rejected in favour of something more explicit. So, instead of *Somsamai Village*, whose definiteness presupposes recognition, a speaker in any doubt might formulate the place as indefinite (*a village upstream of here*), or refer to it less succinctly with a complex description (*the village with the steel bridge, on the way here*). But these informationally more secure options run a different kind of risk, since despite some doubt about the referential adequacy of the selected linguistic formulation (e.g. a proper name), the speaker might at the same time also not be confident that the recipient **won't** achieve reference from the recognitional linguistic expression being supplied (e.g. a simple place name). If the speaker's formulation supposes too little common ground, there is a danger of **over-telling**, being more explicit than was actually necessary, thereby failing to heed a social-affiliational imperative (Schegloff 2007a). In other words, if a speaker formulates his or her utterance so as to put on-record (Brown and Levinson 1987) a presumption that the addressee does not possess the requisite knowledge to achieve reference – e.g. that they wouldn't recognize a particular place name on first mention – the risk is socially distancing the interlocutor by treating them as not sharing knowledge with the speaker (Enfield 2006).

So, while the risk of under-telling is informational/referential failure, the risk of over-telling is epistemic misjudgment and social/affiliational failure. The insecure-reference utterances present the speaker, or better, composite utterance designer, with a dilemma of constraint satisfaction. The Lao speakers' S-points give them a halfway solution: maintain the symbolic, focal, linguistic component of the utterance as it should be, assuming informational calibration, and add a pinch of extra information (here, the spatial information indexed by the gesture) in a relatively low-cost, off-record, non-intrusive, potentially pre-emptive, compensatory form. Or as a slogan: risk under-telling in the speech, risk over-telling in the gesture. For affiliative reasons, speakers will prefer to err on the side of under-telling (or, 'over-supposing', as Schegloff 2007a puts it) rather than over-telling. When teetering in the middle, a well-placed S-point is an informational safety net.

4.1.2.3 *Why S-points are formally well fitted to insecure-reference contexts*

Think about the kind of composite utterance which is home to the S-point. This is a move whose purpose is to converge with the speaker on referential coordination when the adequacy of the linguistic

formulation alone is not guaranteed. As laid out in the previous section, the S-point component of the composite utterance functions to provide a bit of extra information just in case, but where this information should not be perceived as even necessary, since this active attention to a perceived informational risk of under-telling must be balanced against the desire to avoid a social-affiliational risk of over-telling. Here, as we saw for B-points, above, form fits function. The reduced physical movement involved in the articulation of S-points is iconic-indexical of 'not primary assertion' (cf. Levy and Fowler 2000), via a Gricean pragmatic maxim of quantity (i.e. given that the option of doing the bigger gesture is available yet not selected; Grice 1989, Levinson 2000). More importantly here, the speaker's formulation of the hand movement as casual and small is a natural way to overtly downplay the need for its even being there. The S-point's casual, reduced, fast articulation means that the locational information it provides is vulnerable to not being easily interpreted by the recipient. This vulnerability is itself part of the signal. Displayed vulnerability of informativeness is neatly fitted for the function illustrated in the above examples. It may be expected, given the naturalness or non-arbitrariness of these form-function mappings, that these findings will more or less generalize across cultural settings. The speaker indicates by their chosen manner of formulation that the information available in the hand movement is not especially important, as if to say 'you probably don't need this extra information, because you probably know what/where I mean, so I'm not going to too much trouble to make it available to you, since it's probably not going to matter if you don't receive it'. In this sense, the addition of locational information by the S-point gesture is less on-record, and a possible threat to face is minimized.

4.2 A broader consideration of 'secondary pointing' phenomena

If Lao speakers' B-points represent an adult form of proto-linguistic gesture-focal moves with a conventional sign component in the form of speech more or less tacked on, the S-points, by contrast, have an inverted structure. In the S-point composite utterance, speech is primary. The hand gesture component retreats into the finer information structure of the move. We may compare these S-points to types of gesture in a broader category of composite utterances reported in other studies in which the pointing component is backgrounded or otherwise secondary.

Bavelas *et al.* (1992) describe a type of gesture called citing, one of a number of gestures they refer to as 'interactive', since these gestures function to manage the interpersonal/interactional context of speech. This type of gesture is familiar from everyday life, for instance in scientific discussion. A speaker proffers a small pointing gesture, formally identical to the Lao

S-point, in the direction of someone present, while offering a comment which makes reference to something this person has previously or elsewhere said independently. (Imagine you raise the topic of relativity while Einstein is in the room; chances are you will accompany this with a little point or a nod in his direction.) Like a Lao S-point, this form of citation also deals with social-affiliational considerations rather than focal informational content. The example works because of a socially consequential notion in the cultural environment of scientists (with its analogue in every imaginable domain of social expertise), that different individuals have differential rights to claim authority, according to their real or rightful area of knowledge. The problem is, if you make a statement unprompted – i.e. in ‘first position’ in conversation or other discourse – you are thereby making an implicit claim of having primary authority to make that statement (Heritage and Raymond 2005). And if you happen not to be the one in the room with the collectively acknowledged highest authority to say it, you will have to negotiate the mismatch between your own (actual) secondary authority and the (implicit) claim of primary authority via the first position placement of your contribution.

Like the Lao speakers’ S-points, these small ‘citing’ pointing gestures provide a neat solution to the dilemma. A wave of the index finger by the speaker in the direction of the present higher authority is sufficient to explicitly defeat any implicature of primary authority arising from the speaker’s first-position placement of the move. The common link with the insecure-reference environments in the Lao study is that the small pointing gesture is performed in the service of heading off a possible problem at the social-affiliational level. This type of gesture is socially regulated. Its function is not primarily to achieve informational alignment, or even some kind of structural alignment in the interactional sequence. Its function is to deal with interpersonal affiliation. We see again the speaker’s systematic attention to satisfying a social-affiliational constraint, as distinct from an effort simply to achieve referential convergence.⁸

Another type of secondary pointing gesture is observed among the ‘abstract deictic gestures’ identified by McNeill and colleagues (McNeill 1992, 2003, McNeill *et al.* 1993) in data from American English speakers. Abstract deictic gestures are defined formally by their orientation to ‘empty space’.⁹ The

⁸ Dealing with face is a computational problem for speakers, given that face is linked to social intelligence as well as emotion, both central to human cognition (Goffman 1959, Brown and Levinson 1987). If the present analysis is correct, then interestingly, the social imperative outranks the informational imperative when the two come into conflict. A question for further research is whether this ranking is universal.

⁹ The definition of abstract pointing in McNeill *et al.* (1993) includes pointing at empty space in narrative, metanarrative, and paranarrative levels. McNeill (1992: 173) has a slightly narrower

chunks of space indicated by such gestures are temporally endowed with representational value for the purpose of discourse (McNeill 1992: 113ff., 171, 199–200; cf. Chapters 5–7 in Part II, below). We are interested here in a subset of these, which are used for higher-level discourse-organizational tasks, at metanarrative and paranarrative levels. (These may be collectively referred to as supernarrative.) Like the Lao S-points, supernarrative abstract pointing gestures are reduced and restricted in form (McNeill 1992: 199–200). They function to track both concrete and abstract components of a discourse at informationally backgrounded levels. Furthermore, McNeill and colleagues note, in their observations of English and Georgian conversations, that such pointing gestures tend to occur in moves which make the initial few references to a newly introduced entity in a discourse (McNeill 1992: 215–217, McNeill *et al.* 1993: 15–17). This is analogous to the function of S-points in the Lao data. At a moment of potential insecurity in converging on a referent, abstract supernarrative pointing gestures anchor a referent in a gesturally created space, and S-points anchor the referent in the physical environment surrounding the speaker and the addressee.

The common function of these various kinds of secondary pointing gesture – Lao S-points, Bavelas *et al.*'s citing gestures, McNeill *et al.*'s supernarrative gestures – is that unlike primary pointing gestures (B-points) they function at levels of organization which are subordinate to the focal information of the overall composite utterance (i.e. as expressed by the conventional sign component in the speech). Here's the omega end of the pointing spectrum. The two major types of gesture identified here are not just smaller versus bigger. They are qualitatively distinct. Secondary pointing gestures are embedded in composite utterances in complex, subtle, and dependent ways, and have an entirely different life to their proto-linguistic B-point counterparts.

McNeill (1992) raises the issue of the acquisition of children's ability to perform supernarrative gestures. Concrete pointing is the first of all types of gesture acquired (Scaife and Bruner 1975, Bates *et al.* 1975, McNeill 1992: 24, 300; cf. Liszkowski 2006), coming between nine and twelve months of age, significantly earlier than language. By contrast, the use of abstract deictic gestures with metanarrative function 'increases from age 8' (McNeill 1992: 322), while abstract deictic gestures with paranarrative function are not acquired until around age eleven, very late in developmental terms (McNeill 1992: 326). These late-acquired supernarrative functions of gesture presuppose great sophistication in handling information structure in composite

definition, which excludes narrative level pointing that utilizes an iconically established spatial layout.

utterances both internal to the move, and at a higher, inter-move level. McNeill argues that the lateness of acquisition of supernarrative abstract pointing is in line with the late acquisition of clearly distinct narrative levels in the organization of discourse more generally (McNeill 1992: 319). ‘The lateness of abstract pointing is particularly striking given that concrete pointing is the first of all gestures to emerge; abstract pointing is, however, one of the last’ (McNeill 1992: 319–321). The present account – from alpha to omega – fits this picture.

4.3 Distributing informational and affiliational tasks in the formulation of composite utterances

Any cognitive scientific interest in complex external representations in human communication must incorporate the full range of constraints and resources for communicative action in face-to-face interaction.¹⁰ The present investigation raises implications for our understanding of how humans represent meaning in face-to-face interaction, and the kinds of constraints and affordances our habitual mode of representation – enchronic, audience-designed, and fully multimodal – presents us with.

For solving the challenging computational problem of making one’s communicative intentions recognizable to others, people have a range of resources at their disposal. While the largely conventional sign medium of spoken lexico-syntax is linear and digital, other semiotic resources allow additional information to be represented simultaneously with speech. In composite utterances, made up of lexico-syntactic structure, prosody, hand movements, facial expressions, etc. in concert, the conventional sign component (i.e. the lexico-syntax) has a special role, being the most ‘on-record’ component. If we want parts of our overall message to be less on-record, then it is in our interest to keep them out of the lexico-syntax. The solution is to employ semiotic resources which are relatively unobtrusive. A good example is intonation contour, which can be varied independent of lexico-syntax (cf. De Ruiter *et al.* 2006). Directly analogous to one of the S-point functions discussed in this chapter is the try-marker (Sacks and Schegloff 1979/2007). Suppose you want to say *Yesterday I saw John*, but you’re not entirely sure if your addressee will recognize who you mean by *John*. You can ‘try-mark’ the

¹⁰ Some traditions in research on meaning in cognitive science and linguistics have all but ignored the formal complexity of composite utterances, focusing on conventional linguistic semantic structures which underlie lexical (and sometimes constructional) resources. On the other hand, both anthropologists and cognitive scientists have looked seriously at multimodality in perception and cognition where external artefacts are involved (Goody 1977, Larkin and Simon 1987, Norman 1991, Glenberg and Langston 1992, Hutchins 1995, Zhang 1997, Oestermeier and Hesse 2000; cf. Chapter 6, below).

proper name by overlaying it with rising intonation (making it sound like a question). This has the effect of ‘asking’ the hearer to supply confirmation that she understands who is intended, while allowing you to maintain a grammatical formulation which presupposes recognition. A: *Yesterday I saw John?* B: *Uh-huh* (indicating recognition, go-ahead). A: *He was (etc . . .)*. On paper, the formulation ‘Yesterday I saw John’ is done in such a way as to display an assumption that the addressee should know who is meant by ‘John’. The pragmatic modulation achieved by the try-marker does not compromise the informationally optimistic and therefore socially affiliative framing of the utterance’s lexico-syntactic component.¹¹ The semiotic structure of try-marking directly parallels the Lao S-points. In both cases, information is incorporated in the composite utterance in a form which is manipulable independent of, and simultaneously with, the lexico-syntax. Like the Lao S-point, the try-marker is used in initial mentions of recognitionals (i.e. proper names). The common mechanism is that where the addition of certain information might be a threat to face (since it betrays the perceived possibility of a current lack of common ground; Enfield 2006), it is done in such a way that it does not cause the speaker to have to reformulate a more optimistically phrased lexico-syntactic structure.

4.4 Concluding remark

This chapter has described two ways in which speech and hand-pointing are combined in composite utterances. The two kinds of pointing can be related to fundamental properties of pointing as a basic mode of human communicative action (Tomasello *et al.* 2005, Tomasello 2006, Liszkowski 2006). The B-point (big in form) is hardly different functionally to the pointing gestures performed by prelinguistic infants. It is primary for comprehending the basic meaning of the composite location-focus utterance, and in some cases may inform all by itself. Typically, added speech is merely supportive/elaborative of the B-point gesture, and if anything is dispensable it is the speech. S-points (small in form) are of a different order. They are more subtle, more dependent, more hidden in the information structure of the composite utterance. While a B-point is doing the primary work of the utterance, with speech playing a supporting role, an S-point adds a backgrounded modifier to a move in which speech is central. A speaker’s addition of an S-point to a spoken utterance is a low-cost, low-risk communicative action, which might save against a potentially high social-interpersonal cost, i.e. the undesirable

¹¹ Along the same lines, Hayashi (2005) describes the try-marker as a technique for preserving the integrity of the syntax in the best interests of conversational progressivity (Stivers and Robinson 2006).

situation of being on record as having underestimated common ground (Enfield 2006). It is a way of striking a middle ground when two preferences conflict – a socially-driven preference against over-telling versus an informationally-driven preference against under-telling. While the B-point is an openly foregrounded informative action, the S-point is a backgrounded, subtle, pre-emptive strike in the ongoing struggle to maintain intersubjectivity.

P A R T I I

Illustrative components of moves

Language, after all, is our one way of communicating the kinds of things that do not lend themselves to drawing.

Leonard Bloomfield, 1933

5 Modelling

Is it even always an advantage to replace an indistinct picture by a sharp one? Isn't the indistinct one often exactly what we need?

Ludwig Wittgenstein, 1953

This chapter considers composite utterances whose goal is to describe the forms and functions of complex artefacts – in this case, traditional fish traps made and used in Lao-speaking villages. In these utterances, hand movements are in the foreground, modelling and depicting, bearing a significant load in conveying meaning at the utterance level. The data presented in this chapter address a key issue in the semiotics of hand gestures, namely the question of whether non-linguistic components of composite utterances show linear segmentation and combinatoric dependencies. The data discussed below show that the hand movement components of a sequence of composite utterances can involve linear segmentation and combinatoric dependency. This is not to say that gestures are static, arbitrary signs (*pace* McNeill 2005: 75) – they are not.¹ It is perhaps on terminological grounds that one might disagree with a claim that the data discussed in this chapter reveal a combinatorics to sequences of gesture, for example, in the sense that Hasan (1968: 18–24) insists that textual cohesion is ‘not structure’ (cf. Halliday and Hasan 1976). But my aim here is to draw attention to commonalities between semiotic relations within the move and semiotic relations across moves. There are relations of dependency among gestures in enchronically adjacent moves, where the interpretation of one sign is directly dependent on where it stands with reference to another sign. This draws on what is common to all compositionality in human communicative practice, whether it be by means of classical ‘rule-governed’ syntax, hierarchical patterns of phrase structure, mere co-occurrence, or discontinuous dependency.

¹ McNeill’s (2005) arguments against static views of meaning are valuable because they draw attention to the generally problematic nature of a Saussurean view of meaning for **any** semiotic system, not just gesture or speech (Kockelman 2005). Having come to reject a static view of meaning for gesture, it is not necessary to still defend it for speech. One aim of this book is to show that what is required for the proper analysis of gesture turns out to be required for language too (see Chapter 1, above).

The data presented in this chapter demonstrate that co-speech gestures can and do show combinatoric principles and linear-segmented organization. The claims I want to make are that (a) a given gesture may be structurally related, both in form and function, to a gesture in a neighbouring move; (b) information may be segmented and supplied linearly with gestures (and not simply as a direct reflection of such linear-segmented structuring in the spoken code); and (c) multiple gestures can occur simultaneously, where the gestures play significantly different roles (both semantic and pragmatic) within the composite utterance underway.

5.1 A symmetry–dominance construction

This chapter documents a two-phase routine called a symmetry–dominance construction. Phase 1 is a two-handed symmetrical gesture; in the subsequent Phase 2, one hand holds in position (representing given/topical/backgrounded information from Phase 1), while one hand executes a new gesture (representing new/focal/foregrounded information). This is identical to the phenomenon described in sign language as buoys (Liddell 2003) or residue (Mathur 2002), and akin to what McNeill (2005) calls gestural layering. This construction shows features of linear segmentation and combinatoric structure that arise from both the affordances and the constraints of the manual/visuospatial modality.

The conventional sign system of spoken language is profoundly linearized, due to severe constraints on the number of semiotic dimensions along which speech sounds may simultaneously represent information about referents.² (This concerns relative limitations in the capacity of speech sounds to bear iconic and indexical meaning; see below.) Speech phonology facilitates the construction of contrastive lexemes, and patterns of morphosyntax allow linear strings of these lexemes to be interpreted hierarchically, yielding complex novel meanings in phrases and sentences (Hockett 1960, Thorpe 1972, Lyons 1972, Levelt 1989, *inter alia*). Here are two reasons why gestural representations are less subject to linearization and hierarchical organization. First, due to their manual, three-dimensional, visuospatial modality, gestures can directly and simultaneously represent a rich variety of information about referents. Included here are fine analogue distinctions in three-dimensional space, thanks to the extraordinary spatial (orientational, axial, topological) flexibility of the arm and hand (Talmy 2006).³ Second, non-conventional

² Speech can of course simultaneously convey a good deal of information about the signer rather than the referent (e.g. age, sex, socioeconomic background, state of mind etc.).

³ Talmy (2006) masterfully distinguishes some thirty independent semiotic dimensions of hand movements. I disagree, however, with his characterization of these dimensions as reflecting a

gestures are supported online by speech, a highly conventional code, and are therefore not required to carry the full propositional load in composite utterances. They are free to do what they do best, namely represent information analogically, iconically, and indexically.⁴ Gesture's manual/visuospatial modality has rich semiotic affordances, and furthermore gesture is freed from propositional responsibility as a result of fundamental support that it receives from the finely structured symbolic representational system that we call speech. But gesture is not an unlimited resource for the depiction of just any meaning, of any level of complexity, at any moment. Like speech, hand movements reach representational limits. When it is no longer physically possible to represent the full meaning desired at a given moment, one is forced to segment information and distribute its representation across time. This chapter documents a systematic mechanism for the combinatoric organization of two-handed and one-handed gestures, where linearly sequenced gestures in adjacent moves are structurally dependent.

In using the term symmetry–dominance construction, I borrow and adapt the terms symmetry and dominance from research on the phonological structure of lexical signs in sign language (Battison 1978, Engberg-Pedersen 1994, Miller 1994, Emmorey 2002, Sandler 2002). The symmetry condition in sign language phonology states that 'if both hands of a sign move independently during its articulation, then both hands must be specified for the same hand shape and the same movement (whether performed simultaneously or in alternation), and the specifications for orientation must be either symmetrical or identical' (Sandler 2002: 6). The dominance condition specifies that 'if the hands of a two-handed sign do not share the same specification for hand shape (i.e., they are different), then one hand must be passive while the active hand articulates the movement, and the specification of the passive hand shape is restricted to be one of a small set (6 shapes)' (Sandler 2002: 6). In this chapter, I use these terms in a different but related sense. In a symmetry phase of co-speech gesture, the two hands are symmetrical or identical in shape, position, and movement (or they enact a two-handed action). In a dominance phase, one hand is passive (i.e. is held still) while the other hand articulates some representational gesture. The non-dominant hand inherits its form from a gesture performed in a previous utterance, representing

domain of expressivity unique to sign language, since speakers of spoken languages also have hands and also use them in conveying meaning in composite utterances.

⁴ This description of co-speech gestures as relatively free contrasts with phrasing to the opposite effect by Goldin-Meadow *et al.* (1996), i.e. that hand movements are semiotically 'handcuffed' when occurring with speech, and 'liberated' in sign language of the Deaf, where they are no longer interpreted in connection with co-occurring vocalization (cf. Stokoe 1960, Klima and Bellugi 1979, Emmorey 2002). It seems to me that the heavy constraints of grammaticization could be more aptly thought of as enslaving rather than liberating.

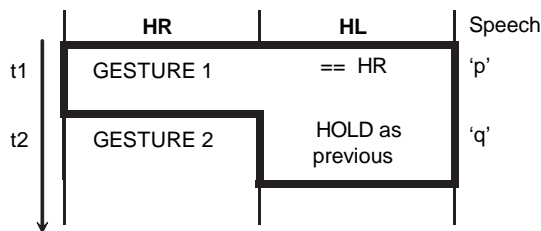


Figure 5.1 The 'half-T' shape (in bold line) of a symmetry–dominance construction.

background information introduced in that utterance, while the dominant hand represents new/focal information. Perhaps the most significant point of theoretical interest concerning the relation between gesture and the structuring and production of speech is that the gesture performed by the non-dominant hand does not mirror the spoken utterance semantically or pragmatically (cf. McNeill 1992: 26–29).⁵

In the schematic Figure 5.1, the symmetry–dominance construction appears as a dark-bordered 'half-T' shape. The columns headed HR and HL represent a speaker's right hand and left hand, respectively.⁶ The passage of time runs downward from row to row. Each row represents a new gesture by either hand, where a distinction between one gesture and another is marked by a thick black bordering line.

Thus, in Figure 5.1, the right hand HR is performing GESTURE 1 at t1, and GESTURE 2 at t2. Now look at the left hand HL. At t1, HL is performing GESTURE 1 in concert with HR. GESTURE 1 is a two-handed gesture. The structure of such a gesture will typically be symmetrical, such that the form of HL can be in some way derived from the form of HR. This is represented in Figure 5.1 using a double equals sign: '=='. Accordingly, there is no thick black line separating the cells HR at t1 and HL at t1, because the two cells do not represent different gestures. However, at t2, as HR is employed in making a second gesture, HL maintains a hold in the hand shape and spatial position it acquired at the end of t1 (marked 'HOLD as previous' in Figure 5.1). Accordingly, there is no thick black line separating the two cells HL at t1 and HL at t2. A symmetrical gesture followed by a hold of the non-dominant hand and a new gesture by the dominant hand results thus in a

⁵ In any case, principles of 'synchrony' in meaning such as those proposed by McNeill (1992: 26–29) are not rules that are followed in formulating composite utterances, rather they are interpretive strategies, which work to impose convergent interpretations upon co-occurring signs, even when there is some degree of incompatibility (cf. Figure 1.3, Chapter 1).

⁶ The position of right and left hand are as they would be for the viewer of a video recorded from in front of the speaker. The speaker's right hand is at the left for the viewer/reader.

visible ‘half-T’ shape in the diagram. A key component of the construction is the hold of one hand (HL in this case) across adjacent moves (t1 and t2).

The two sequential phases of this construction have a defined structural relationship in both form and meaning. This type of construction is typical in descriptions of complex physical artefacts, where a speaker refers to multiple features of a three-dimensional object. Typically, speakers start an artefact description by introducing the overall shape and size of the artefact. For a speaker to begin by introducing the larger structure of which all else relevant can be said to be a part is well motivated in terms of information supply in discourse. Having established the overall size and shape of an artefact, a speaker then has a background context for making finer specifications. The dimensions of the fish traps discussed in examples below make them suitable for speakers to use largish symmetrical gestures when initially establishing their basic whole form. This first step should provide an anchor, or frame, for subsequent information supply in two structured conceptual domains: discourse and three-dimensional space. The syntactic dependency between the two phases of this construction operates in these two domains simultaneously, ensuring coherence of both the organization of discourse and the representation of spatial structure.

5.2 Examples

We now turn to description of four sequences, two descriptions each of two types of basket trap for fish: the *sòdn5* and the *tum4 thoong2*. The examples are from interviews in which speakers were asked to describe types of locally produced fish-trapping mechanisms, including nets, basket traps, drop-door traps, filtering gear, box enclosures, fences, and spearing equipment (see Claridge *et al.* 1997). In describing fish-trapping artefacts, speakers give complex performances involving a range of iconic and indexical hand gestures, creating virtual objects in shared space. A critical feature of the symmetry–dominance construction in these performances is the semiotic relation between the non-dominant hand (the hand that holds its position and shape established in the symmetry phase) and the dominant hand (the hand that makes the new focal gesture). The non-dominant hand is used extensively as a counterpart and anchor for the dominant hand, providing semiotic support. This support is twofold. First, the non-dominant hand provides a stable spatial reference point (or ground) facilitating the depiction of complex three-dimensional spatial representations by the dominant hand. Second, deployment of the non-dominant hand has a discourse pragmatic or prosodic function, signalling that certain background information continues to be relevant to what is being said. Meanwhile, the dominant hand directly represents information foregrounded in the current speech (cf. Sandler 2002,

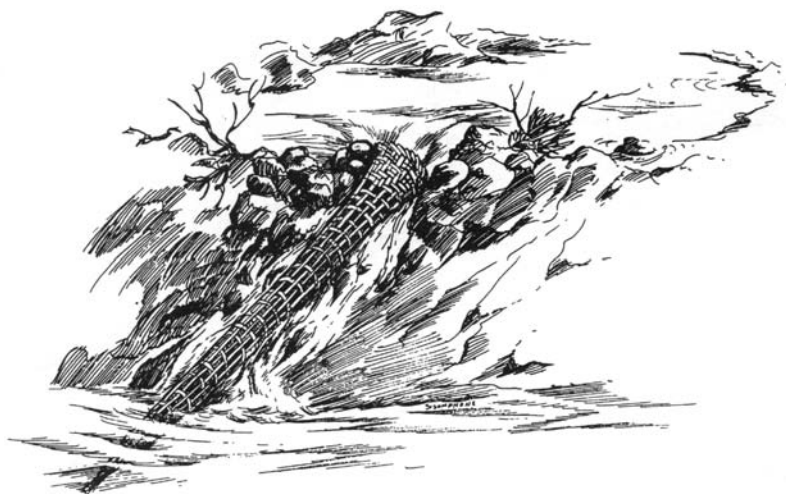


Figure 5.2 The *sòn5*, a cone-shaped woven ‘wedge’ basket fish trap. (Illustration by Sisomphone Hansai [Claridge *et al.* 1997: 54].)

Mathur 2002, Liddell 2003: ch. 8). These functions of the non-dominant hand have consequences not only for relative spatial deployment of the two hands, but also for their respective temporal deployment.

5.2.1 Example 1 – *sòn5*

The *sòn5* is a cone-shaped woven wedge basket trap, usually a metre or so long. It tapers from a fluted mouth (between six and twelve inches across) down to a tight end. It is placed in flowing water, such that the water enters the open end, causing any fish swimming into it to be trapped inside and unable to get back out, due mainly to the fish’s inability to turn around in a tight space.

5.2.1.1 *Speaker 1* The speaker begins his description of the *sòn5* as follows:

- (1) *khan2 sòn5 lèw4 khaw3-*
 if *sòn5 PRF 3PL.B-*
khaw3 hêt1 [vee3 thaang2 paak5 man2 naø]
 3PL.B make fluted direction mouth 3.B TPC.PERIPH
 ‘As for the *sòn5*, well they- they make it [fluted at the mouth]_{Fig. 5.3a.}’

The first recognizable gesture in this sequence occurs when the speaker comes to the word *vee3* ‘fluted’. The gesture is a two-handed symmetrical representation of the fluted opening of the trap.



Figure 5.3 (a) ‘As for the *sòòn5*, well they- they make it fluted at the mouth’; (b) ‘(They) make (it) small . . .’; (c) ‘. . . going in’.

Illustration	HR	HL	Speech
Fig. 5.3a	depicting trap body and mouth	== HR	‘...they make it fluted at the mouth.’
Figs. 5.3b–c	depicting fish going into trap	HOLD as previous	‘(They) make (it) small, going in...’

Figure 5.4 The ‘half-T’ in the diagram shows a symmetry–dominance construction (Figures 5.3a–5.3c).

In the next move, going into a dominance phase, the left hand is held in position, while the right hand represents the passage of a fish into the mouth and body of the trap. Figures 5.3b and 5.3c represent the beginning and end of a single movement.

- (2) [hêt1 nòjə-nòj4] [khaw5 paj3]
 make RDP.A-small enter go
 ‘(They) [make (it) small]_{Fig. 5.3b}, [going in]_{Fig. 5.3c}.’

This simple example of the symmetry–dominance construction is diagrammed in Figure 5.4.

The example illustrates the typical function of the symmetry–dominance construction. First, a two-handed gesture establishes a larger structure that forms a spatial frame of reference for subsequent specifications about the artefact. When the left hand holds in the position established during the symmetry phase, it continues to represent the referent of that phase – i.e. the larger structure – even though by itself it constitutes literally only half of the original signifier. The held left hand indexically signifies what had been

represented by the left and right hands together during the symmetry phase. It provides a spatial frame of reference for the elaborating right hand gestures in the dominance phase. Note that at the phase depicted in Figure 5.3a, the focus of interest in the current utterance is the fluted mouth of the trap (note speaker's eye-gaze and head-orientation). The rest of the trap is implied, but its precise length and orientation are not specified. When the speaker releases his right hand from the symmetrical gesture shown in Figure 5.3a and begins the dominance-phase gesture in Figure 5.3b, his left arm is the only remaining signifier of the trap as a whole, and it determines the orientation of the body of the trap. The right-hand gesture in Figures 5.3b and 5.3c specifies a precise orientation of the body of the trap, namely running from the space at the front and centre of the speaker's body back to his left elbow, at the left side of his lower torso. This location is then presupposed in subsequent moves in this sequence.

Here we see a combinatoric relation of coherence between gestures in adjacent moves, where the subsequent gesture is formally and semantically subordinate to the first. The first gesture sets up a discourse and spatial frame for following gestures, using full available manual semiotic resources – i.e. both hands. This full image persists thanks to a metonymic relation between the degraded (exactly halved) signifier (i.e. just one hand) in Figures 5.3b and 5.3c, and the fuller conceptual image required for current spatial and discourse cohesion.

The speaker's next statement concerns the very tight weave at the small end of the trap. The gesture he chooses to accompany his speech in representing this is a two-handed enacting gesture, as if weaving. The gesture is performed at the point in space where the small end of the trap is now established to be located at the left side of the speaker's lower torso, where his left elbow was in Figure 5.3b.

- (3) *lèw4 khaw3 [saan3 nòj-òj4] maa2*
 PRF 3PL.B weave RDP.A-small come
 '... and they [weave (it) small]_{Fig. 5.5a} coming (back here) ...'

Here, the speaker has interrupted a dominance-phase because he needs both hands for his next gesture, in this case an enactment of a two-handed action, weaving. The next gesture, in Figure 5.5b, leads back to the same symmetrical gesture shown in Figure 5.3a above.

- (4) *laø khaw3 maa2 hêt1 [vee3 thaang2 paak5 man2 nòq1]*
 PRF 3PL.B come make fluted direction mouth 3B QPLR.
 AGREE
 '... and they come (and) make (it) [fluted at the mouth, right]_{Fig. 5.5b}.'

This symmetrical gesture provides the basis for entry into another dominance-phase with the left hand held, this time for longer than the sequence

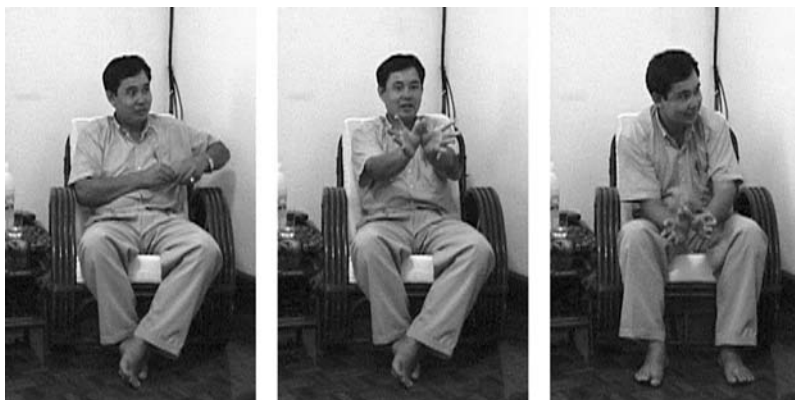


Figure 5.5 (a) ‘... and they weave (it) small, coming (back here) ...’;
 (b) ‘... and they come (and) make (it) fluted at the mouth, right’;
 (c) ‘And (they) place it in the rice fields, also.’

shown in Figures 5.3b and 5.3c. The speaker now wants to talk about how the trap is actually used. Still in symmetrical phase, the speaker moves his body forward, maintaining the symmetrical and held two-hand gesture, shifting it forward and down, as if placing the trap down (Figure 5.5c).

- (5) *lèkaø qaw3 paj3 [saj1 thaang2 naa2 khùu2 kan3]*
 C.LNK take go put direction paddy same COLL
 ‘And (they) [place it in the rice fields, also]Fig. 5.5c.’

The left hand now holds, as the speaker enters a dominance phase, using his right hand to represent a fish that gets caught in the trap. The following three lines and accompanying illustrations (Figures 5.6a–c) depict the entry of the fish into the trap and its getting stuck inside. Note that the speaker’s gaze follows the right hand all the way in.

- (6) *vêlaa2 [paa3 siø long2] baat5-niø*
 Time fish IRR descend THZR
 ‘Now, when a [fish is going to go down]Fig. 5.6a (into it) ...’

- (7) *paa3 siø long2 laø man2 [khaw5 paj3]*
 fish IRR descend PRF 3.B enter go
laø paj3 [sup1] juu1 han5
 PRF go insert be.at there
 ‘... (when) a fish is going to go down (into it), it [goes in]Fig. 5.6b and [is inserted]Fig. 5.6c there ...’



Figure 5.6 (a) ‘Now, when a fish is going to go down (into it) . . .’;
(b) ‘. . . (when) a fish is going to go down (into it) . . .’; (c) ‘. . . it goes in
and is inserted there . . .’

Then, the speaker states that the fish cannot get back out. His right hand, moving quickly forward and away from his body, represents a fish coming out (Figure 5.7a), while his speech provides the clausal negation (cf. Engle 1998).

- (8) *laø man2 [khùun2 bòø daj4] naø*
PRF 3.B return NEG can TPC.PERIPH
‘. . . and it [can’t go back]Fig. 5.7a.’

Before finishing the sequence, the speaker reiterates the fate of the fish – stuck inside the trap – depicting it once again with a right hand gesture, using a repeated jamming-in movement of his right hand (Figure 5.7b).

- (9) *khaw5 paj3 laø paj3 [ñat1 juul naj2 hanø] lèq1*
enter go PRF go jam be.at inside DEM.DIST FAC.PRF
‘(It) goes in and (it) gets [jammed in there]Fig. 5.7b.’

Figure 5.8 shows a diagram of the sequence depicted in Figures 5.5c through 5.7b, where the heavy-bordered area depicting the symmetry–dominance cycle forms a ‘half-T’ shape.

The speaker is now apparently satisfied that no further elaboration is necessary. As he makes a closing comment as to the yield of this type of trap (not great), his two hands go to rest as he settles his body back in his chair (Figure 5.7c).

- (10) *kaø bòø daj4 laaj3 qanø-niø daj4 nòj5*
T.LNK NEG get much MC.INAN-DEM get little
‘(You) don’t get many (fish), (with) this one, (you) get a little Fig. 5.7c.’



Figure 5.7 (a) ‘... and it can’t go back’; (b) ‘(It) goes in and (it) gets jammed in there’; (c) ‘(You) don’t get many (fish), (with) this one, (you) get a little.’

Illustration	HR	HL	Speech
Fig. 5.5c	Depicting trap move forward and low, as if being placed	== HR	'And (they) place it in the rice fields, also.'
Figs. 5.6a–c	depicting fish swimming into trap	HOLD as previous	'Now, when a fish is going to go down (into it)...(when) a fish is going to go down (into it), it goes in and is inserted there'
Fig. 5.7a	depicting fish coming out of trap, hold outside trap		'and it can't get back.'
Fig. 5.7b	depicting fish going inside trap, with repeated movement of jamming, holding inside trap		'(It) goes in and (it) gets jammed in there.'

Figure 5.8 Diagram of the symmetry–dominance construction shown in Figures 5.5c–5.7b.

The physical act of assuming a resting posture has a prosodic function (in the sense of Sandler 1993, 2002). Putting the hands to rest can signify a discourse boundary, a ‘putting to rest’ of the current topic. This is achieved semiotically by an indexical relation between a speaker’s commitment to a communicative project, on the one hand, and the degree to which he is expending physical effort in carrying out that project, on the other. The latter information is typically directly perceptible (Levy and Fowler 2000). When the speaker’s hands assume a resting posture after a stretch of gesture in discourse, he displays that he is momentarily finished (although the display is not necessarily recognized by addressees as significant in the same way as the intentionally communicative actions preceding it; Kendon 1978: 305). By contrast, the message conveyed by the **non-resting** posture



Figure 5.9 (a) ‘(It’s) small . . .’; (b) ‘. . . and tapered . . .’; (c) ‘. . . fluted like this, the mouth of it’.

of a held gesture – as in a dominance phase – is one of ‘**not** putting to rest’ a current topic. The message conveyed by the poised non-dominant hand throughout the dominance-phase illustrated in Figure 5.8 is that the speaker is not yet done with the topic of the *sòdɔn5* fish trap. Keeping the non-dominant hand in play requires a persistent expenditure of effort, and this is a perceptible signal that the information represented is relevant for current discourse purposes (Levy and Fowler 2000).

5.2.1.2 Speaker 2 Our second example of a description of the *sòdɔn5* has compelling similarities to the example just discussed. This speaker begins with a description of the overall shape of the conical trap, making a two-handed 3D-tracing gesture running his hands symmetrically forward and upward along the imagined body of the trap, from tight-woven base to fluted opening (Figures 5.9a–c). Eye-gaze follows gesture all the way.

- (11) [nòjɔ- nòjɔ4] [suajɔ- suaj3] [pên3 vee3 qòòk5 cang1- sii4,
RDP.A- small RDP.A- taper be fluted exit thus
paak5 man2]
mouth 3.B

‘(It’s) [small]_{Fig. 5.9a} and [tapered]_{Fig. 5.9b}, [fluted like this, the mouth of it]_{Fig. 5.9c}.’



Figure 5.10 (a) ‘And it tapers down like the arm, like this’; (b) ‘Smallish ones- If- if a big one it comes down it gets inserted’; (c) ‘. . . and it can’t turn around’.

The symmetrical gesture phase depicted in Figures 5.9a through 5.9c ends with a hold (shown in Figure 5.9c). This is followed by a long dominance phase, in which the speaker’s right hand plays the non-dominant role, maintaining more or less the hand shape established in Figure 5.9c. As in the previous example (cf. Figures 5.3a through 5.3c above), the switch into the dominance phase anchors the body of the trap in the non-dominant arm, thus orienting it to the side of the speaker’s body rather than to the centre (where it seems to be in the first instance; cf. Figures 5.9a and 5.9b).

The speaker now elaborates on the shape of the trap, using the shape of his non-dominant arm as a point of reference, as he ‘shapes’ it with his dominant (in this case left) hand (Figure 5.10a).

- (12) *laø man2 suaj3 long2 maa2 [khùù2 khèèn3*
 PRF 3.B taper descend come like arm
cang1-sii4 lèø]
 thus PRF
 ‘And it tapers down [like the arm, like this]_{Fig. 5.10a.}’

The speaker then turns his attention to the manner in which a fish becomes stuck in the trap. As the next example shows, after a moment of disfluency, he begins explaining the fate of a fish that enters the mouth of the trap. As he utters the key word *sup1* ‘inserted’, he uses his left

(currently dominant) hand to represent the fish moving down and inserting itself inside the trap (Figure 5.10b; cf. Figures 5.3b through 5.3c and 5.6a through 5.6c):

- (13) *qanø nòjø-nòòj4- khan2- khan2 toø ñaj1*
 MC.INAN RDP.A-small if if MC.ANIM big
man2 long2 maa2 laø man2 [sup1]
 3.B descend come PRF 3.B insert
 ‘Smallish ones- If- if a big one it comes down it [gets inserted]Fig. 5.10b.’

He remarks that the fish cannot then turn around. As he says this, he works his left elbow back and forth as if trying to extract his left hand (i.e. the fish), without success (Figure 5.10c).

- (14) *laø man2 [quaj3] bðø daj4*
 PRF 3.B turn.around NEG can
 ‘and it can’t [turn around]Fig. 5.10c.’

He reiterates the consequence of the fish’s inability to turn around, namely its inability to get out of the trap. This is accompanied by movement of the left hand forward, upward, along, and finally away from his still held non-dominant (right) arm. The movement represents passage of the fish out of the trap (shown in Figure 5.11a, where the left hand is blurred due to its fast movement away from the speaker). This outward passage is negated in the speech (cf. Figure 5.7a above).

- (15) *quaj3 bðø daj4 thùù3 vaa1 [khùn5] bðø daj4*
 turn.around NEG can regard COMP ascend NEG can
 ‘(It) can’t turn around, that is (it) can’t [go up]Fig. 5.11a.’

The sequence illustrated in Figures 5.9a through 5.11a is another typical symmetry–dominance construction, very similar to the one illustrated in Figures 5.3a through 5.3c above. In the symmetry phase (Figures 5.9a through 5.9c), a two-handed gesture establishes size, shape, and overall structure of the artefact. The dominance phase begins when one hand holds still while the other proceeds to make new, focal gestures in subsequent moves (beginning in Figure 5.10a). The information that was focal during the symmetry phase is background/topic information for the dominance phase. The non-dominant hand indexically, metonymically, represents the larger structure established in the symmetry phase, even though it constitutes literally only half of the original signifier.

The speaker now seems to be considering bringing his description to an end. He says ‘That one’, referring to the trap, as a kind of punctuation mark. This is an idiomatic way of wrapping up a small stretch of discourse in Lao (cf. English *That’s it*). Suggesting imminent closure of a discourse



Figure 5.12 (a) ‘A *khêng1* fish, why can’t it . . .’; (b) ‘. . . turn around?’; (c) ‘It’s stuck (on) its gills.’

The speaker explains that the fish’s gills are what prevent it from moving backward. As he does so, he points with his left hand to his own neck, i.e. where his gills would be were he a fish (Figure 5.12c).

- (18) *man2 [khaa2 ngeep4 man2]*
 3.B stuck gill 3.B
 ‘It’s [stuck (on) its gills]_{Fig. 5.12c.}’

In the short sequence shown in Figures 5.12a through 5.12c, the speaker has made a dramatic change in perspective, but he has nevertheless kept in play the structure of the trap, as introduced at the beginning of the sequence (Figures 5.9a through 5.9c). This is achieved by continued visibility of the non-dominant (here, right) hand, which, while lowered in Figure 5.11b, maintains the slightly clawed 5-hand shape that it inherited from the original symmetrical gesture which began this sequence (Figure 5.9c). The speaker now switches back from this parenthetical stretch in character perspective and adopts once again the earlier token space or observer perspective. This is done by simply raising the non-dominant hand (without having to adjust its shape) and returning it to its position immediately prior to the aside (i.e. as in Figure 5.11a). The speaker thus brings the temporarily backgrounded trap back on stage, where it now may serve as a frame for further elaboration of how it is that a fish gets stuck in this type of trap. The speaker’s following comments are accompanied by a gesture whereby the dominant (left) hand represents a fish coming down into the trap (Figures 5.13a and 5.13b).



Figure 5.13 (a) ‘It goes down . . .’; (b) ‘. . . head first’; (c) ‘And they- they can’t get out, (of) that one’; (d) ‘(That’s the) *sòdn5*.’

(19) *man2 qaw3 [hua3 long2]*

3.B take head descend

‘It [goes down head first]Fig. 5.13a, 5.13b.’ (lit. ‘It descends with the head.’)

The non-dominant hand alone is enough here to representationally invoke the entire trap. It represents only one half of the signified object, but by a metonymic relation to the fuller representation shown in Figure 5.9c above, it signifies the whole of it. In Figures 5.13a and 5.13b, this is necessary, because the speaker does not have three hands (i.e. two to represent the trap as in Figure 5.9c, and one to represent the fish going in).

This example demonstrates the way in which representational constraints of the manual/visuospatial modality force segmentation and linearization of information, across adjacent moves. The speaker first establishes the larger structure with full use of manual/visuospatial resources (Figures 5.9a through 5.9c). Subsequently, half of these resources are kept in play, thereby serving to remind the addressee (and probably the speaker himself) of that larger structure. The other half of the resources used in the first stage are then freed for representational elaboration. For this to work, two things are presupposed. Interpreters must be able to retrieve reference to larger structures when these structures are only partially represented by available signifiers (i.e. interpreters must be able to comprehend metonymic relations). Second, and relatedly, distinct gesture structures must be viewable as semiotically dependent across enchronic sequences of moves. The half-signifier structure in the second stage (Figure 5.10a onward in this diagram) would be hard to comprehend unless you had already been exposed to the first stage of fuller representation (Figures 5.9a through 5.9c). During the dominance phase, the non-dominant hand is combinatorically subordinate to the two-handed gesture in the preceding symmetry phase. Here, gesture performs a kind of simultaneous construction (cf. the same in sign language; Engberg-Pedersen

1994, Miller 1994, Vermeerbergen *et al.* 2007), as part of a larger interclausal structure that strives to conform to principles governing the supply of information over time. Here, we see enchronic interleaving of composite utterance components being utilized to build structure above the level of the move. It is because the trap was more fully represented by the symmetrical gesture that opened this sequence (Figures 5.9a through 5.9c) that it could later be evoked by the presence of one hand only (Figure 5.10a onward).

The speaker now moves toward completion of his description. Along with the spoken utterance in (20), he performs a gesture that is a repeat of Figures 5.13a and 5.13b (gesture not shown here).

- (20) *qaw3 too3 long2 man2 tòòng4 qaw3 hua3 long2*
 take body descend 3.B must take head descend
 ‘(To) get (its) body down, it must take (its) head down (first).’

The speaker once again lets both hands down, yet still does not abandon the hand shape assumed by the non-dominant hand, as shown in Figure 5.13c.

- (21) *laø man2- man2 khùn5 bòø daj4, qanø-nanø*
 PRF 3.B 3.B ascend NEG can MC.INAN-TPC.NONPROX
 ‘And they- they can’t get out, (of) that one.’

The speaker is now finished with his account of the *sòòn5*. His closure of this sequence involves a repeat of the name of the trap, again as a kind of punctuation mark, as he finally disengages the non-dominant (right) hand from its persistent representational role, using it now to scratch his left arm, thus causing there to be no longer any perceptible trace of the referent that was established in Figure 5.9c and that had persisted up until this point.

- (22) *sòòn5*
sòòn5
 ‘(That’s the) *sòòn5* Fig. 5.13d.’

The diagram in Figure 5.14 depicts the entire sequence, explicated in Figures 5.9a through 5.13d and accompanying linguistic examples, above. The symmetry–dominance construction is observable as a tall ‘half-T’ shape.

5.2.2 Example 2 – *tum4 thoong2*

The next two cases are descriptions of the *tum4 thoong2*, a rare and old-fashioned type of large basket-woven fish trap, illustrated in Figure 5.15.

It is a type of *tum4* ‘upright basket trap’. The *tum4 thoong2* is pear-shaped, with an opening of about six inches in diameter in the base. Around the rim of this opening, ‘tusks’ stick up into the interior of the trap, allowing fish to enter the opening easily from beneath but exit with difficulty. A long stake is tied fast to the trap body, such that the stake protrudes from both above and

Illustration	HR	HL	Speech
Figs. 5.9a–c	== HR	3D-trace, merging into depiction of mouth of trap	'(It's) small and tapered, fluted like this, the mouth of it.'
Fig. 5.10a	HOLD as previous	running hand along HR arm to demonstrate tapering	'And it tapers down like the arm, like this.'
Fig. 5.10b		depiction of fish going into trap	'Smallish ones- If- if a big one comes down it gets inserted'
Fig. 5.10c		depiction of fish trying to turn	'and it can't turn around.'
Fig. 5.11a		depiction of fish going up out of trap	'(It) can't turn around, that is (it) can't go up.'
Fig. 5.11b	(lowers)	TO REST	'That one.'
Figs. 5.12a–b	(lifts)	moving head to left then to right, taking character perspective (self as fish)	'A <i>khêng1</i> fish, why can't it turn around?'
Fig. 5.12c		pointing to own neck (fish's gills)	'It's stuck (on) its gills.'
Figs. 5.13a–b (not shown)		depiction of fish going into trap	'It goes down head first.'
		depiction of fish going into trap	'(To) get (its) body down, it must take (its) head down (first).'
Fig. 5.13c	(lowers)	TO REST	'And they- they can't get out, (of) that one.'
Fig. 5.13d	TO REST	as previous	'(That's the) <i>sònn5</i> .'

Figure 5.14 Diagram of the symmetry–dominance construction shown in Figures 5.9a–5.13d.

underneath the trap body. The stake is then wedged into a riverbed, such that the trap is fully submerged, but suspended above the riverbed. The entry through the base of the trap is poised six inches or so above the riverbed. Bait is placed inside the trap, around the entry hole, enticing fish to enter the trap from below. After fish are inside the trap, they are free to swim around in the large, open body of the trap interior. They are unlikely to escape, thanks to the uninviting inward-pointing 'tusks' surrounding the entry hole.

5.2.2.1 Speaker 1 In the first part of this sequence, the speaker discusses the type of location where a *tum4 thoong2* would normally be placed and elaborates on the manner of fixing it to the riverbed. We begin here when he starts to discuss the shape and structure of the trap and the manner in which fish find themselves caught in it. The speaker makes a complex symmetrical gesture, first a 3D-trace of the body of the trap (Figures 5.16a and 5.16b), followed by a two-handed depiction of the circular hole at the base of the trap (each hand representing one half of the hole), with upturned fingers standing for the 'tusks' around the entry hole (Figures 5.16c and 5.16d).

- (23) *tum4* [bak2]-[ñaj1] *laø* [hêt1] [ngaa2 saj1
 upright.basket.trap very-big PRF make tusk put
thaang2 lum1
 direction below
 '(It's a) [big]_{Fig. 5.16a} [tum]_{Fig. 5.16b}, and (they) [make]_{Fig. 5.16c} [tusks at the base]_{Fig. 5.16d}.'



Figure 5.15 Resident of a village near Vientiane, Laos, prepares to place a *tum4 thoong2* in the Mekong River. The trap's base features an open hole with inward/upward pointing 'tusks'.

In Figure 5.16d, the speaker is depicting the 'tusks' at the trap base that allow fish easily up and in, but not easily out. This is a symmetrical pose, forming the entry to a dominance phase. The speaker now holds his left hand in



Figure 5.16 (a) ‘(It’s a) big . . .’; (b) ‘. . . *tum4* . . .’; (c) ‘. . . and (they) make . . .’; (d) ‘. . . tusks at the base’.

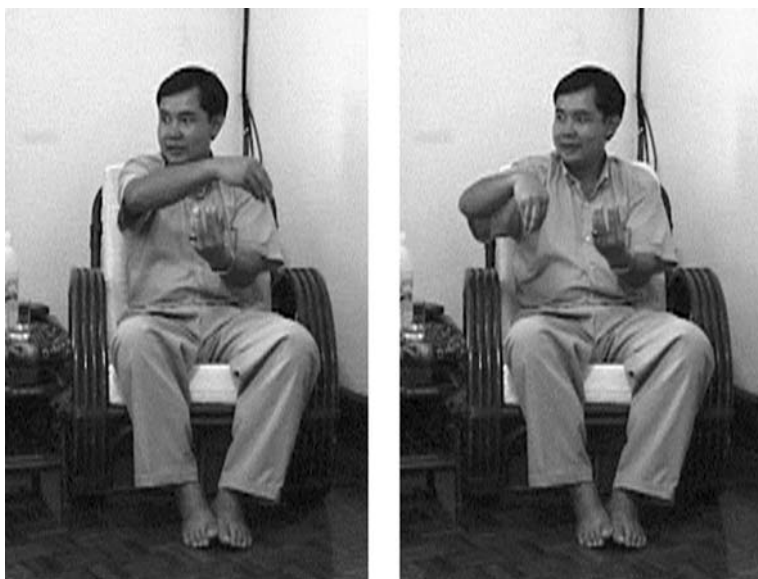


Figure 5.17 (a) ‘(They) take toasted rice . . .’; (b) ‘. . . (and) put (it) there’

place, representing the topical element of the opening in the trap’s base, while using his right hand to introduce new and focal information (pragmatically matching the speech), enacting with a counter-clockwise circular motion the placement of bait in the area around the hole on the interior floor of the trap.

- (24) *qaw3 [khaw5-cii1] [saj1 han5]*
 take rice-toast put DEM.DIST
 ‘(They) take [toasted rice]_{Fig. 5.17a} (and) [put] (it) [there]_{Fig. 5.17b.}’

Illustration	HR	HL	Speech
Figs. 5.16a–b	3D-trace, creating large object in centre front	==HR	(It's a) big <i>tum4</i> ,
Figs. 5.16c–d	depicting fingers up imitating 'tusks' at base of trap	==HR	and (they) make tusks at the base.'
Figs. 5.17a–b	enacting, circling, imitating placement of bait around hole	HOLD as previous	'(They) take toasted rice (and) put (it) there.'

Figure 5.18 Diagram of the symmetry–dominance construction shown in Figures 5.16a–5.17b.



Figure 5.19 (a) '(We) feed (it) down along um the stick, and then (we) go (and)-'; (b) 'So, this is- is- is- is . . .'; (c) ' . . . the earth'; (d) 'Estimate that it come (and stay) at- at- at- at around this level'.

Figure 5.18 shows this symmetry–dominance construction in diagrammatic form (cf. Figures 5.4, 5.8, and 5.14 above).

The speaker then elaborates on a number of things, which for reasons of space I will just summarize here. He talks about placing the bait, in preparation for lowering the trap into the water and fixing it to the riverbed by means of a stake fixed on to it. He elaborates on how the stake is tied to the trap and how the stake is measured such that the trap will be positioned at the right height in the water.

Now the speaker is going to describe the manner in which he lowers the trap (now tied on to a long stake) down into the water. In this composite utterance (Figure 5.19a), the two-handed gesture is not literally symmetrical, but the two hands do signify a single idea (by enactment of the activity of lowering the trap down into the water).

- (25) *pòòn1 long2 paj3 taam3 qanø [lak2] hanø lèèw4 paj3-*
 feed.in descend go follow HES stake TPC.DIST PRF go
 '(We) feed (it) down along um the [stick], and then (we) go (and)-Fig. 5.19a.'

Then, the right hand is held, and with his left hand now in a B-hand shape (i.e. flat palm and extended fingers, fingers together), palm-down, fingers pointing forward, the speaker depicts the riverbed (Figure 5.19b). Why does he choose to use his left hand for this dominance-phase move? He may be anticipating an immediate reversal of role of his left hand, from its status as the one active hand in a phase immediately following a two-handed gesture (Figure 5.19b) to its status as the non-dominant hand, constituting a spatial and discourse-topical point of reference for a long series of right-hand gestures to come (beginning in Figure 5.19c):

- (26) *kaø [qanø-nii4 mèn1 mèn1 mèn1 mèn1] [khii5-din3]*
 T.LNK MC.INAN-DEM be be be be earth
 ‘So, [this is- is- is- is]_{Fig. 5.19b} [the earth]_{Fig. 5.19c}.’

He says in clarification that by ‘earth’ he means the earth at the riverbed. Then he wants to say that the trap is lowered until there is a gap of six inches or so between the trap’s base and the riverbed. In representing this, he makes a held gesture with his right hand (like the left hand, in a B-hand shape, palm-down, fingers pointing forward), poised six inches or so above the already-held left hand (Figure 5.19d). The right hand depicts the trap’s base while the left hand depicts the riverbed. Note the eye gaze on the speaker’s own gesture, here and throughout this section.

- (27) *kaq2 vaa1 haj5 man2 maa2 juul juul juul*
 estimate COMP cause 3.B come be.at be.at be.at
[juul ñaam2 nii4]
 be.at level DEM
 ‘Estimate that it come (and stay) at- at- at- [at around this level]_{Fig. 5.19d}.’

He repeats that the trap should stay at around the level depicted and then elaborates on the question of the space between the trap base and the river floor. As he makes the statement in (28) he moves his right hand directly down, maintaining its posture (as in Figure 5.19d), until it meets the left hand, depicting the non-desired situation in which the trap base touches the riverbed (Figure 5.20a).

- (28) *khan2 [cam1 laaj3] man2 kaø khaw5 bòø daj4*
 if close/tight very 3.B T.LNK enter NEG can
 ‘If (it) is [very close]_{Fig. 5.20a} (to the riverbed), then it (i.e. the fish) can’t get in.’

He is then asked to confirm that the fish are meant to enter the trap through the hole in the base. He replies ‘yes’ and elaborates. With his held left hand still representing the earth floor of the river, he depicts with his right hand a fish swimming along the riverbed (Figure 5.20b), in through



Figure 5.20 (a) ‘If (it) is very close (to the riverbed), then it [i.e. the fish] can’t get in’; (b) ‘It- It down along the bed . . .’; (c) ‘. . . (and) it sees (it and) . . .’; (d) ‘. . . then it goes in’.

the base of the trap (Figure 5.20c), and up into the interior of the trap (Figure 5.20d).

- (29) *man2* [*man2*] *long2* *maa2* *nam2* *khii5-din3*
 3.B 3.B descend come with earth
 [*man2*] *hên3*, *man2* *kaø* [*khaw5*] *paj3*
 3.B see 3.B T.LNK enter go
 ‘It- [It]_{Fig. 5.20b} comes down along the bed (and) [it]_{Fig. 5.20c} sees (it and) then it [goes in]_{Fig. 5.20d}.’

He then adds that the fish will eat the bait, before uttering a discourse marker *qee5* ‘uh-huh’, as if closing down this section of the discourse. In accordance with a closure of topic, he lets his right hand go to rest. Note that he does not yet rest his left hand – the hand that is currently representing a backgrounded but anchoring referent, the river floor below the trap – but keeps it in play (exactly as in the previous example; cf. Figure 5.11b and subsequent figures).

Then, he is asked to verify that the fish cannot get out. In reply, with left hand still held, he represents with his right hand a fish swimming around inside the expansive interior of the trap, using a high circulating counter-clockwise sweep of his right hand (Figures 5.21b and 5.21c).

- (30) *laø* *man2* *kaø* [*huu4* *thaang2*
 PRF 3.B T.LNK know direction
lòj2] *dêj2* *baat5-niø* *man2* *ñaj1*
 float FAC.NEWS THZR 3.B big
dêj2- *bòø* *mèn1* *too3-nòj4* *dêj2*
 FAC.NEWS NEG be MC.ANIM-small FAC.NEWS
 ‘And they [have a place to float around]_{Fig. 5.21b, 5.21c} you know, it’s big you know. It’s not a small (fish), you know.’

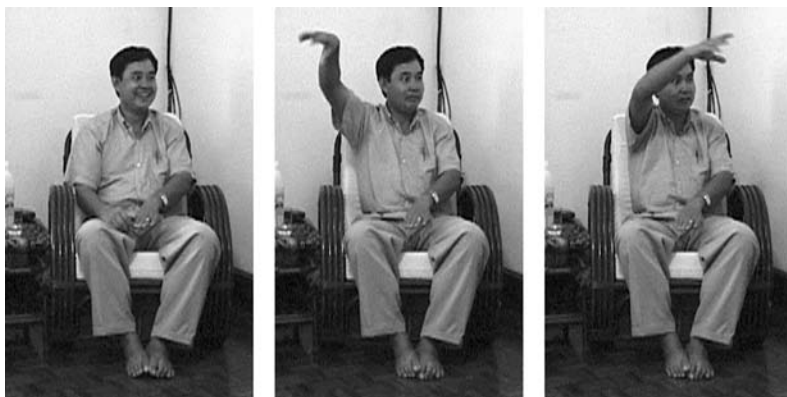


Figure 5.21 (a) Speaker with right hand at rest, left hand still poised (representing backgrounded referent); (b) ‘And they have a place to float around you know . . .’; (c) ‘. . . it’s big you know. It’s not a small one, you know.’

The speaker’s left hand has remained in play since Figure 5.19b, held in the same position, and assuming a non-dominant anchoring role. It has provided a persistently perceptible signifier of a backgrounded feature of the discourse (the earth at the riverbed), which has served as a structural anchor, both in discourse information structure and in spatial arrangement of the relevant entities and their subparts. The sequence depicted in Figures 5.19a through 5.21c can be diagrammatically represented as in Figure 5.22.

The first two rows of Figure 5.22 show a symmetry–dominance construction in which the left hand plays the dominant role. In subsequent rows, the left hand takes on the non-dominant role, and the right hand elaborates throughout the rest of the sequence.

Continuing his description, the speaker now states that once set, the trap should be left for some time. One later goes back to visit it and collect any fish that have become trapped inside. He states that when one goes to visit the trap, one first gently releases it from the stake by untying it. As he says this, he makes a one-handed gesture with the right hand, enacting untying the trap from the stake, yet with the left hand still poised in the position shown in Figures 5.19b through 5.21c. (For reasons of space, this sequence is not shown here.) Finally, he frees his left hand from this long-held position, in order to facilitate a new two-handed gesture, an enactment of pulling up the

Illustration	HR	HL	Speech
Fig. 5.19a	enacting, lowering the trap using the pole	== H1	'(We) feed (it) down along the um the stick and then (we) go (and)-'
Fig. 5.19b-c	HOLD as previous	depicting river bed	'So, this is- is- is- is the earth.'
	touches HL, indexically, including rub back and forth	HOLD as previous	'-um- the earth at the bottom of the water.'
Fig. 5.19d	depicting base of trap, displaying size of gap		'Estimate that it come (and stay) at- at- at- at around this level.'
Fig. 5.20a	depicting base of trap, showing lack of gap		'If (it) is very close (to the riverbed), then it can't get in.'
Figs. 5.20b-c	depicting fish, coming to base of trap, up in through mouth, and into centre of inside trap		'It goes in underneath?' 'Yes.'
Fig. 5.21a	TO REST, falls to lap		'It- it comes down along the bed (and) it sees (it and) then it goes in. (And) it eats the toasted rice in there.'
Figs. 5.21b-c	Depicting fish, swirling motion of fish floating around inside trap		'Yeah.'
			'And then it can't get out?'
			'And they have a place to float around you know, it's big you know. It's not a small one, you know.'

Figure 5.22 Diagram of the symmetry–dominance construction shown in Figures 5.19a–5.21c.

trap out of the water, shown in Figure 5.23a.

- (31) *man2 laø khòðj1 saaw2* [khùn5
 3.B PRF gentle manipulate.long.stick.as.tool ascend
maa2 naø]
 come TPC.PERIPH
 'And (we) gently [pull it up]_{Fig. 5.23a} (using the stick).'

He continues this action, repeating 'gently pull it up, gently pull it up'. Then, he describes what is happening to the fish while the trap is being lifted up through the water. Because the base hole is a fixed opening, it is possible for a fish to escape if it happens to be right above the hole when the trap is raised. The speaker's gestures resume the role of representing structures associated with the body of the trap. We first observe a new symmetrical gesture, with palms facing down, elbows up, representing the inside surface of the trap (Figure 5.23b).

- (32) [*phùin4 man2 kuang4 dêj2]*
 floor 3.B broad FAC.NEWS
 '[The base of it is broad, you know]_{Fig. 5.23b}'



Figure 5.23 (a) ‘And (we) gently pull it up (using the stick)’; (b) ‘The base of it is broad, you know.’



Figure 5.24 (a) Symmetrical gesture, prefiguring reference to circular hole in trap base; (b) ‘But at the mouth . . .’; (c) ‘. . . it’s roundish’.

This is followed by another symmetrical gesture, which prefigures a reference to the hole in the base of the trap, where the two hands depict the diameter and partial shape of a circular opening (Figure 5.24a).

This symmetrical gesture is the first phase of a new symmetry–dominance construction. In Figure 5.24a, the speaker represents the hole in the trap’s base with his two hands. Next, the left hand is held where it is (now assuming a non-dominant role), while the right hand assumes a dominant role, making a circular motion (counter-clockwise) tracing the roundish shape of the hole (Figures 5.24b and 5.24c).

- (33) *tèè1 vaa1 [bòòn1 paak5 niø] man2 monø-mon2*
 but COMP place mouth TPC 3.B RDP.A-round
 ‘But [at the mouth]Fig. 5.24b, 5.24c, it’s roundish.’

In Figures 5.24b and 5.24c, the non-dominant hand plays the subordinate role of representing a backgrounded, topical structure that the focus of the current utterance (i.e. what the dominant hand represents) is about. The non-dominant hand iconically represents only a fragment of the circular opening that it is intended to signify. It metonymically evokes the fuller image of that referent already established in the symmetry phase that preceded it (Figure 5.24a). The non-dominant hand provides an anchoring spatial point of reference for the dominant hand’s movement, as well as providing continued perceptible signification of topical information (i.e. the larger structure that the focus of the utterance is about). The brief symmetry–dominance construction illustrated in Figures 5.24a through 5.24c may be diagrammed as in Figure 5.25.

The speaker now breaks the dominance phase depicted in Figures 5.24b–c, because he needs both hands for his next move. The subsequent gesture is symmetrical, depicting once again the ‘tusks’ that stick up from the base hole into the interior of the trap (Figure 5.26a). The speech follows on directly from example (33).

- (34) *laø hêt1 [ngaa2] khùn5 juu1 nòòj5-nùng1 dêø*
 PRF make tusk ascend be.at a.little FAC.FILLIN
 ‘And (they) make [tusks]Fig. 5.26a coming up a little.’

The speaker now holds his left hand in place, moving into another dominance phase. The left hand continues to represent the full symmetrical structure established in Figure 5.26a, namely the inward-tusked opening to the trap. Again, it does this by virtue of having been part of a fuller representation at an earlier phase. The held left hand provides a spatial and topical anchor for what the right hand now does, which is to depict the straight downward motion of a hypothetical fish that happens to be above the opening

Illustration	HR	HL	Speech
Fig. 5.24a	3D-trace/depicting, showing hole at base of trap	== HR	
Figs. 5.24b–c	index-finger trace, circling, shape of hole	HOLD as previous	‘But at the mouth, it’s roundish’

Figure 5.25 Diagram of the symmetry–dominance construction shown in Figures 5.24a–5.24c.



Figure 5.26 (a) ‘And (they) make tusks coming up a little’; (b) ‘But if any one (of them) . . .’; (c) ‘. . . were it to strike the opening, then it’d certainly get out’; (d) ‘But they don’t tend to strike (it).’

as the trap is raised, thus allowing the fish to escape, as depicted in Figures 5.26b and 5.26c.

- (35) *tèl1 vaa1 khan2 [too3-daj3] man2 thùuk5 [pòòng1*
 but COMP if MC.ANIM-INDEF 3.B strike opening
hanø lèø man2 kaø siø qòòk5 paj3 hanø
 TPC.DIST PRF 3.B T.LNK IRR exit go TPC.DIST
lèø
 FAC.PRF

‘But if [any one]_{Fig. 5.26b} (of them), were it to strike [the opening]_{Fig. 5.26c}, then it’d certainly get out.’

With left hand still held, the speaker returns the right hand to recreate the symmetry that began this sequence (Figure 5.26d; cf. Figure 5.26a).

- (36) *tèl1 vaa1 [man2 bòø suu1 thùuk5 hanø]*
 but COMP 3.B NEG tend strike TPC.DIST
 ‘But they [don’t tend to strike (it)]_{Fig. 5.26d}.’

The sequence depicted in Figures 5.26a through 5.26d is diagrammed in Figure 5.27.

The rest of the sequence (not shown here) involves a series of two-handed gestures enacting the removal of the trap from the water, and of the fish from the trap, finishing with the speaker resting his hands.

5.2.2.2 Speaker 2 The fourth and final segment to be discussed in this chapter is from a second speaker’s description of the *tum4 thoong2*. This segment features close parallels to the sequence just examined, beginning with a symmetry–dominance construction almost identical to that depicted in Figure 5.18, above. A two-handed depiction of the tusked opening in the base

Illustration	HR	HL	Speech
Fig. 5.26a	depicting, fingers up imitating 'tusks' at base of trap	== HR	'And (they) make tusks coming up a little.'
Figs. 5.26b–c	depicting fish passing down through hole at base	HOLD as previous	'But if any one (of them), were it to strike the opening, then it'd certainly get out.'
Fig. 5.26d	depicting, fingers up imitating 'tusks' at base of trap	== HR	'But they don't tend to strike (it).'

Figure 5.27 Diagram of the symmetry–dominance construction shown in Figures 5.26a–5.26d.

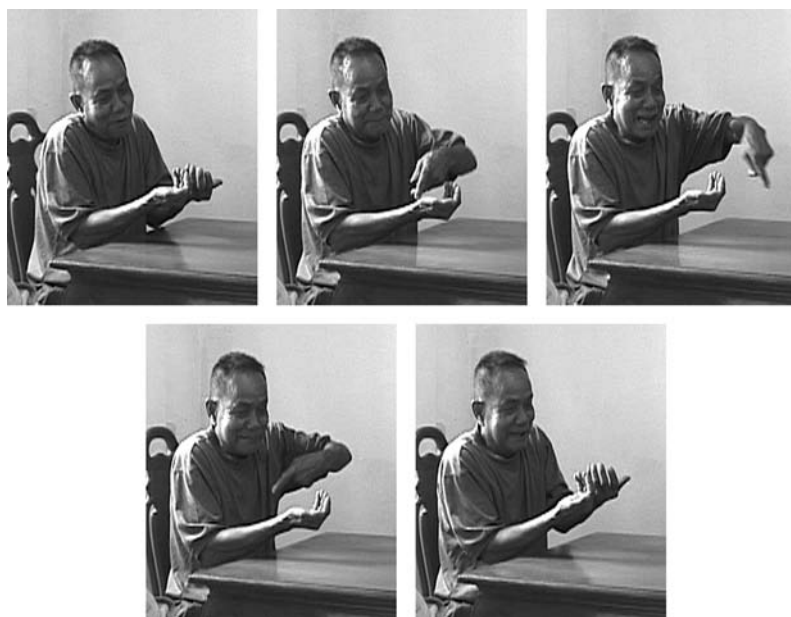


Figure 5.28 (a) 'That's why it's woven, going *vok vok vok* like this'; (b) 'Here it's all a platform', a; (c) 'Here it's all a platform', b; (d) 'Here it's all a platform', c; (e) 'Here it's all a platform', d.

of the trap is followed by a left-hand hold and a series of elaborating gestures performed by the right hand, assuming the left hand as a spatial and topical anchor.

After the speaker has made a few preliminary remarks, he is asked whether bait is used and where the bait is placed. He replies that it is placed around the tusks on the floor of the trap's interior. He explains that the trap is woven so as to have a platform-like floor surrounding the tusked hole at the base. He begins with a palms-up symmetrical gesture depicting the tusked nature of the hole at the trap's base (cf. the almost identical gesture by a different speaker

Illustration	HR	HL	Speech
Fig. 5.28a	depicting, tusks at base of trap, palms out HOLD as previous	== HR	'That's why it's woven, going vok vok vok like this.'
Figs. 5.28b–d		deictic, circling, pointing out where bait lies	'Here it's all a platform.'
Fig. 5.28e		== HR	

Figure 5.29 Diagram of the symmetry–dominance construction shown in Figures 5.28a–5.28e.

shown in Figures 5.16c and 5.16d above). The tusks stick up and inside the trap, which means that whatever is placed on the platform around it (in the interior of the trap) will not fall through the hole (Figure 5.28a).

- (37) *kòòl man2 [saan3 pên3 vok1 vok1 vok1] cang1-sii4*
 that's.why 3.B weave as IDEO IDEO IDEO thus
 'That's why it's [woven, going vok vok vok]_{Fig. 5.28a} like this.'

Now, in a replica of the symmetry–dominance construction illustrated in Figures 5.17a and 5.17b, above, the left hand is held in position, maintaining the job of depicting the tusked base hole, while the other hand takes the dominant role, indicating the area of the trap floor, around the hole in the trap's base, where bait is placed. This is done with a counter-clockwise-circling index-finger point (Figures 5.28b through 5.28d), the left hand then returning to its former symmetry position (Figure 5.28e; cf. Figure 5.28a).

- (38) *nii4 man2 [pên3 saan2 mot2 dêø]*
 DEM 3.B be platform all FAC.FILLIN
 'Here [it's all a platform]_{Fig. 5.28b–5.28e}.'

Once again, the non-dominant hand in Figures 5.28b through 5.28d only literally signifies half of its original referent. It single-handedly depicts a two-handed image, a feat made possible only by an enchronic link in relations of form, meaning, and linear order between gestures in neighbouring moves. The right (here non-dominant) hand in Figures 5.28b through 5.28d maintains a discourse presence, conceptually, of the fuller structure depicted in Figure 5.28a. If the speaker had three hands, he would not have to rely on this technique. In response to the limitations of the modality, he has segmented and linearized the information for the purpose of supplying it to an interlocutor in a coherent way.

This symmetry–dominance construction is diagrammed in Figure 5.29.

After further elaboration (not shown here), the speaker produces a different representation of the tusked opening, this time with palms turned inward, as indicated in Figure 5.30a.



Figure 5.30 (a) Palms-in representation of tusked opening in base of trap; (b) ‘They go up . . .’; (c) ‘... through this one’; (d) Hands to rest, signalling closure.

Illustration	HR	HL	Speech
Fig. 5.30a	depicting, tusks at base of trap, palms in	== HR	‘They go up along/through this one.’
Figs. 5.30b–c	depicting, fish going up into trap opening		
	
Fig. 5.30d	TO REST	TO REST	

Figure 5.31 Diagram of the symmetry–dominance construction shown in Figures 5.30a–5.30d.

The symmetrical gesture shown in Figure 5.30a comprises the first stage of a final example of the symmetry–dominance construction. The left hand is now held still, continuing to represent the opening in the trap’s base, while the right hand depicts new information, namely a fish going up into the interior of the trap through the base (Figures 5.30b and 5.30c).

- (39) *man2 khùn2 [nam2 qanø-nii4]*
 3.B ascend with MC.INAN-DEM
 ‘They go up [along/through this one]_{Fig. 5.30b, 5.30c.}’

After a few further comments (not shown here), the speaker finishes by bringing both hands to full rest, indexing a withdrawal of communicative effort and signaling a close to his contribution (Figure 5.30d).

Figure 5.31 shows the sequence from Figures 5.30a through 5.30d in diagrammatic form.

This closes the descriptive portion of this chapter. Let us take stock of what these examples reveal.

5.3 Semiotics of the symmetry–dominance construction

How does the symmetry–dominance construction convey the meanings it conveys? I note three important mechanisms.

First, there is semiotic unity in symmetry. In the symmetry phase, the fact that the two hands formally match each other in shape, position, timing, and motion indexes a semiotic unity. The perceptible symmetry is a trigger for an interpreter to take the two hands together to express a single idea. They are readily seen as one (and indeed can hardly be seen as separate).

Second, in the dominance phase, the non-dominant hand has a metonymic relation to the two-handed representation in the preceding symmetrical phase. As exemplified above, during the dominance phase, the held hand indexes the full conceptual and spatial structure established in a preceding phase. In these examples, the external representation has no full form, but is imagined at the invitation of the speaker and sustained by whatever means necessary and sufficient. Clearly, partial signification is enough. And note that the second sign (one-handed) stands simultaneously for the first sign (two-handed) and for its referent, in a classical discourse-cohesive structure. It differs from the kinds of reduction we get in, say, pronouns, since in the case of these gestures the later, reduced form is in a causal relation to, and has formal qualities in common with, an earlier, full form. The referential fragment persists in perceptible form across a sequence of adjacent moves, unlike the invisible yet implied form of a referent in semi-activated common ground.

A third key semiotic feature of this construction is its signification of information about discourse structure and a speaker's current communicative contribution. This is indexical of the speaker's internal state, with reference to their communicative intention – i.e. information about the signer, not the referent. This relates to 'grounding of reference in perception' (Levy and Fowler 2000), the notion that co-speech gesture (among other types of signal) is a means of publicly displaying the degree of energy one is expending in order to communicate. Moving the hands as part of one's communicative contribution is a perceptible index that one is indeed making an extra communicative contribution. One's intentional involvement in a communicative project might otherwise be imperceptible, because in essence it inhabits the mental interior and would not be accessible without 'conversion from private to public' (Miller 1951: 3). It follows that to cease such effort is to publicly display that one is no longer, at that moment, devoting such effort to the task of communication. Accordingly, full relaxation of the hands after gesturing during speech is a powerful index

of a closure boundary in discourse.⁷ Examples in the sequences discussed above include Figures 5.7c, 5.13d, and 5.30d.

5.4 Semiotics of the non-dominant hand: an interclausal function of the hold in gesture

A defining component of the symmetry–dominance construction is the gesture hold. One known function of the hold in co-speech gesture is to make online adjustments to speech–gesture synchrony in order to avoid potential transgressions of a phonological synchrony rule (a ‘strong urge to keep gesture and speech together’; McNeill 1992: 26). For example, a gesture whose stroke comes too early may enter a post-stroke hold until the relevant prosodic peak in accompanying speech arrives. These are intra-move functions. The data described in this chapter show that a gesture hold’s scope of function is not confined to the boundaries of a single move. The above examples show that a gesture hold is a kind of enchronic glue, inhabiting a structural position in the temporally ordered supply of information not only within single moves but also across moves in a discourse trajectory. By formally linking a new composite utterance with a preceding one, a gesture hold affords a coherent way to explicitly represent two related ideas at the same time. The information represented by the held non-dominant hand is carried forward from a previous move, relevant to the current move but pragmatically backgrounded. Thus, the gestures simultaneously articulated during a dominance phase each perform distinct pragmatic functions with respect to the speech.⁸

5.5 A remark on sign language

The phenomenon discussed in this chapter may occur in any type of composite utterance in which the hands are involved. This includes not only composites of speech with gesture, but also sign language of the Deaf (see Liddell’s 2003 description of ‘buoys’). Sandler (1993, 2002), for example, discusses the role of the non-dominant hand in lexical signs. The non-dominant hand either copies the dominant hand’s shape, position, and movement,

⁷ See Kendon (1978) for the differential attention that addressees pay to main track gestures as opposed to disattend track actions such as putting the hands at rest.

⁸ This is inconsistent with a proposed pragmatic synchrony rule for gesture: ‘if gestures and speech co-occur they perform the same pragmatic functions’ (McNeill 1992: 29). Clearly, the speech and gesture components of a composite utterance are taken to be mutually relevant, and each points to a single overarching utterance meaning, but within this we see that they have distinct pragmatic functions (foregrounded versus backgrounded, topical versus focal etc.).

or it is a passive and abstract anchor for a novel contribution by the dominant hand.

Sandler (2002) says that while the non-dominant hand is semantically redundant, it contributes prosodic information. In the articulation of lexical signs, the non-dominant hand may (a) mark ‘the boundaries of the prosodic word’ (by finishing off a two-handed sign while the dominant hand has already moved on to begin a new sign), or (b) mark ‘a prosodic constituent’ by ‘spreading’ (holding in anticipation of an oncoming two-handed sign, waiting for the dominant hand to finish its current performance of a one-handed sign; Sandler 2002: 14–17). It is the third of Sandler’s prosodic functions of the non-dominant hand that is comparable to its role in the symmetry–dominance constructions described here. As Sandler puts it, at ‘the discourse level’, articulations of the non-dominant hand may ‘break the bounds of the signs and sentences in which they originate by persisting in the signal to track and background referents and events’ (Sandler 2002: 18; cf. Emmorey and Falgier 1999, Liddell 2003: ch. 8). Mathur (2002), similarly, describes the import of the non-dominant hand in one type of hold in sign language as representing semantic ‘residue’, persistent meaning carried forward from the sign in which it was first introduced (i.e. where ‘some part of a sign is preserved through the articulation of the following signs’; Mathur 2002: 5). This accords with what we observe in the speech-with-gesture composites, above.⁹

5.6 Concluding remark

The kinds of composite utterance described in this chapter demonstrate combinatoric principles above the level of the move. These holds of the significant hand are mortar for moves in their role as bricks for building higher-level discourse structure. Combinatorics is normally discussed in the domain of intra-move relations, that is, with reference to the elaboration of internal structure of moves as units of contribution to discourse. But as we have long known from functionally-oriented approaches to language, the sentence or equivalent is not a suitable cut-off line for the phenomena of interest to a theory of human communication. While we may argue as to whether patterns above the level of the move constitute ‘structure’ in the same way as patterns within the move (Halliday and Hasan 1976), both domains

⁹ Mathur argues that to keep a hand in play with residual meaning is ‘is to allow articulatory ease’: ‘it is easier to keep the H2 [the non-dominant hand] in place until it is needed for the next two-handed sign’ (Mathur 2002: 5). But it is surely **more** effort to keep a hand raised in the air. To rest the hand would be to invite an interference that the information it represents is no longer relevant. This is appropriate at a boundary of discourse closure, which is precisely where we do observe the hands going to rest.

display principles of co-occurrence and sequential ordering that determine how utterances are to be interpreted. This is the sense in which I argue that these sequences of gesture show a kind of inter-move combinatorics.¹⁰

The symmetry–dominance construction is an information-packaging technique based on semiotic principles afforded by the manual/visuospatial modality. The extent to which this is equivalent to a grammatical construction in speech is open to discussion. It has some of the elements. Dominance sequences are structurally dependent on symmetry sequences. As the Lao fish-trap descriptions exemplify, the construction is used when the informational structure to be conveyed is more than can be represented with a single gesture by a single human being (with two arms and one head). Speakers in these circumstances will segment, package, and supply the information in linear order, for essentially the same reasons that linearization occurs in speech. And the relation between the form of the gestures and the order in which they appear will not be random. The result is a gestural combinatorics, building structural dependencies in both the spatial and discourse domains, through the linear-segmented organization of hand movements in enchronic sequences of moves.

¹⁰ To make this claim is not to argue that gestures are semiotically ‘static’ (*pace* McNeill 2005: 75).

6 Diagramming

The artist cannot copy a sunlit lawn, but he can suggest it.

E.H. Gombrich, 1960

Hand movement is a powerful medium for diagramming thought. With hand movements, the ordinary speaker routinely creates complex visual illustrations. These representations are visible and vivid during production and then, like speech, gone from the perceptual field. What is the link between the evanescent diagrammatic practices of gesture and visual representations of the historically enduring kind, such as iconographies, written texts, or printed diagrams? This chapter pursues the idea that the human body is a cognitive artefact both for those who inhabit it and for those who view it (Hutchins 1995, 2006). This is most vividly illustrated by focusing on hand gestures, perhaps the most salient contribution to the rich geography of cognition (Goodwin 2000a) constituted by bodily positioning, gaze, talk, and orientation to both the physical environment and the attentions of co-present others.

Different forms of iconographic or diagrammatic representation can have different cognitive consequences in problem-solving (Larkin and Simon 1987, Zhang 1997, Oestermeier and Hesse 2000). While this suggests implications for the broader study of culturally situated forms of visual representation, research on the role of external artefacts in cognition (e.g. Norman 1991, Hutchins 1995) has yet to connect with the ethnographic study of visual representations in non-literate and/or non-technological settings (but cf., for example, Wassmann 1997, Wilkins 1997, Green 2008 among other works on representation of spatial information). At the same time, the wealth of research on iconographic and other depictive visual representations (e.g. Marshack 1972, Munn 1986 [1973], Dubinskas and Traweek 1984, Layton 1991) has not explored implications for cognitive processing. Certain other types of visual representation – specifically those enduring visual representations which are capable of directly representing speech – have led some to argue for profound cognitive consequences on a historical level (Goody 1977, Ong 1982, Donald 1991, Olson 1994; cf. Street 1984). But despite a mass of research on the written word, we still know little of the role of nonverbal visual forms such as iconographic depictions or diagrams in the

online cognitive processes of those who produce and interpret them (Scaife and Rogers 1996). Existing work on the relation between diagrams and thought has concentrated on graphs, tables, and charts, which represent only a fraction of culture's visual products. We have yet to recognize the extent to which indigenous visual representations, both enduring and non-enduring, may constitute 'technologies of the intellect' (Goody 1968).

I explore these themes with reference to the spontaneous visual representation of abstract kinship relations in ethnogenealogical interviews (cf. Conklin 1969 [1964]). Diagrammatic representation of kinship relations is not only a technical practice within the culture of social science but a revealing cultural and cognitive practice among lay people. While kinship research – a giant among topics in the history of anthropology – has always relied on diagrams, little scholarly attention has been paid to kinship's visual depiction itself as a topic for empirical ethnography (Bouquet 1996; cf. Conklin 1969 [1964]). To what extent are our technical kinship diagrams modern products of the domesticated mind (Goody 1977)? Could their spatial properties just as well emerge from 'savage' conceptions?

In this chapter I present a study of bodily movements made by semi-urban speakers of Lao during informal ethnogenealogical interviews. Video-recorded segments show speakers faced with the communicative challenge of defining features of their own system of kinship and kin terminology. This challenge elicits spontaneous use of semiotically composite representations – sequences of speech in concert with sequences of hand gestures, eye-gaze movements, torso orientations, and other meaningful physical actions. These efforts are visibly designed to be viewed by recipients and, further, are straightforwardly recognized by onlookers to be part of what the speakers are saying. The examples show speakers using hand movements and other bodily movements as tools for diagramming thoughts on a rich three-dimensional virtual sketch space anchored in the body.

6.1 Co-speech gesture as a means of external representation

While signs in different modalities can convey apparently identical information, the employment of one rather than the other can imply different kinds of thinking, including different effects on memory, reasoning, and imagination (Goody 1977, Ong 1982, Larkin and Simon 1987, Glenberg and Langston 1992, Zhang 1997, Goldin-Meadow 1999, 2003a, Kita 2000, Oestermeier and Hesse 2000, Emmorey 2001). In composite utterances of the kind described in this book, different types and sources of information are complementary and co-constitutive of a larger whole message. The composite utterance par excellence involves simultaneous integration of (conventional) speech, (symbolic indexical) gesture, and (non-conventional, iconic-indexical) visual

representations or artefacts. In this way, the weather reporter gives verbal commentary while using a pointer to link his or her speech with marks and symbols on a graphic display. In a less formal setting, I explain to you how to find my house, using my index finger as a link between my verbal commentary and the lines and symbols on a city map. A diagram drawn or printed on paper and an analogous diagram ‘drawn’ by hand movements in mid-air have in common the property of being visually accessible externalized communicative representations. Both are produced by people for the purpose of organizing and conveying information in joint social activities. They can be seen and therefore have spatial properties.

A fundamental difference between drawn or printed diagrams and diagrams sketched by hand movements in mid-air is that the latter are evanescent.¹ One advantage of adding hand movements to speech is that one becomes able to use absolute coordinates in physical space for tracking referents across moves in a spoken discourse (Marslen-Wilson *et al.* 1982: 356, McNeill *et al.* 1993, Liddell 1995, 2000, McNeill 2003). In the absence of visible artefactual diagrams, hand movements are often used as if perceptually accessible artefacts were actually in view. This is how pointing and tracing gestures are usually employed to construct virtual diagrams indexically. We can point to chunks of thin air, assigning referents to stable points in space and inviting our interlocutor to imagine that such a diagram is hanging there. Recent work on sign language also explores this phenomenon (Liddell 2000, 2002, 2003, Emmorey 2001, 2002).

6.2 Semiotics of gesture diagrams

A concrete pointing gesture is an indexical signal involving some vector-projecting articulator (such as gaze or an index finger) oriented toward a physical object or location, in which the gesture is understood to have the communicative purpose of indicating that thing or some aspect of it (or a deferred referent; cf. Quine 1971). (See Part I, this volume.) It is easy to think that a concrete pointing gesture simply links a movement of the hand with a thing in the world, but this is wrong (Wittgenstein 1953: 13–18). All forms of pointing set up a conceptual referent – the concept of what is being pointed to – and in concrete pointing, the conceptual referent happens to be mapped onto the same piece of space occupied by a physically existent referent. What is being indicated may of course not even be an entity, but perhaps a colour

¹ Evanescence is a matter of degree. Sand drawings or drawings on a chalkboard persist as long as they are not rubbed out. Only rarely do these types of representation stand the test of time. An example is the glass encased chalkboard at the School of Physics and Astronomy, University of Nottingham, upon which Einstein wrote lecture notes during a visit in June 1930.

exemplified by some entity (Wittgenstein 1953: 16). Liddell (1995) points out the need to differentiate between the actual, existing physical space and what he calls real space – ‘an individual’s conception of what is physically real in their current, directly perceivable physical environment’ (Liddell 1995: 23). If you indicate an object by pointing at it, you set up a conceptual referent (in your conceived real space), and this conceptual referent is also instantiated by the physical thing. The physical thing thus becomes semiotically activated.

In abstract pointing (McNeill *et al.* 1993), we similarly set up a conceptual referent, but we do so by pointing not to a physical object or location, but to a chunk of space. A chunk of space is a segment of physical space in the environment of a speaker that (1) is persistently delimited only by interlocutors’ actively imagining it to be delimited; (2) is delimited by virtue of having been assigned, by gesture and speech in combination, a signifying relation to some referent active in the discourse; and (3) has stable spatial coordinates after having been established, in general not shifting in absolute space despite subsequent movements of a speaker’s body. Such chunks are no bigger than a soccer ball (not coincidentally, I presume, since this is about the size of something a person could comfortably hold – and behold – in two hands). A chunk of space inherits its characterization from publicly visible performance in interaction (along lines described by Goodwin 2000a, *inter alia*). It is an invisible but stable target, created by people behaving as if a real thing were there, floating in shared space, when in physical reality there is nothing.

Liddell (1995) differentiates between two types of invisible chunks of space used in American Sign Language. Surrogate spaces are ‘full-sized invisible entities’, which have ‘body features, being viewed as present with the signer’ (Liddell 1995: 27–28). In signed languages, this may involve imagining a person physically standing by. Signs that are normally directed at different parts of the body of an interlocutor (e.g. forehead, chin, chest) can be similarly directed to different levels of height in physically empty space when referring to someone who is not present (cf. Taub 2001: 79ff.). A model of a sand castle sketched using co-speech gesture is another example of a surrogate space. Token spaces, by contrast, are ‘featureless’ reference-activated blobs of space ‘sized to fit inside the physical signing space where signs are articulated’ (Liddell 1995: 33). (In fact, they are not restricted to the body-immediate gesture space, as will emerge below.) It is common, for example, for pointing gestures directed toward a speaker’s left and right gesture space, respectively, to correspond with references to two distinct or contrasting ideas in a discourse (McNeill 1992: 173ff.). This is an especially common strategy for pronominal reference in sign language (Liddell 2000, 2003).

For some purposes it is crucial to differentiate between surrogate and token space, but for other purposes it is useful to treat them together simply as cases of semiotically activated temporarily stable delimitations of thin air, and it is

for this reason that I use the more general term **chunk of space**. Space itself is inherently undifferentiated, and the idea that chunks of it can be isolated in the first place is due to our imposed chunking **of** it, no doubt due to the differentiating effect that the human torso has on space (i.e. turning undifferentiated space into gesture space). Gesture space has inherent coordinates in the human body that projects the space. Further, interactional space itself – for example, the general area in and around which interlocutors are situated – may be accorded internal structure according to a range of factors associated with cultural conventions, interactional dynamics, and physical features of the space (Chapter 2, above; cf. Goffman 1963, Kendon 1977, Schefflen 1976, *inter alia*). Space, in the important sense of interactional space, thus cannot unproblematically be said to be undifferentiated. By pointing to empty space, we invite an addressee to imagine that something is there. By looking like you are holding a ball, you can create in your interlocutor's mind the idea of a ball; analogously, by looking like you are pointing at a physical object, you create in your interlocutor's mind the idea of that object (cf. Haviland 2000: 20).

Reference-activated chunks of space can be multiplied and can together form structured complex wholes, such as the diagrammatic illustrations discussed in this chapter and the next, and the 3D 'models' discussed in the [previous chapter](#). The simplest case involves two chunks of space (to the left and right of the gesture space, respectively) referring to two distinct referents being tracked in a discourse, where the two chunks maintain a constant spatial relation to each other. Despite their ethereal nature, chunks of space are treated like physical objects, and the same can be true of whole structured arrays of multiple chunks of space. The reference-activated chunks and their configuration relative to each other show temporal persistence and a certain kind of spatial integrity or cohesion. The configurations can persist over time, as if a complex object were there in the real physical space.

Certain types of conceptual problem are particularly amenable to diagrammatic representation, for example, where one is required simultaneously to maintain reference to multiple individuals and refer to multiple (abstract) relationships between them. The kinship domain provides a good example.

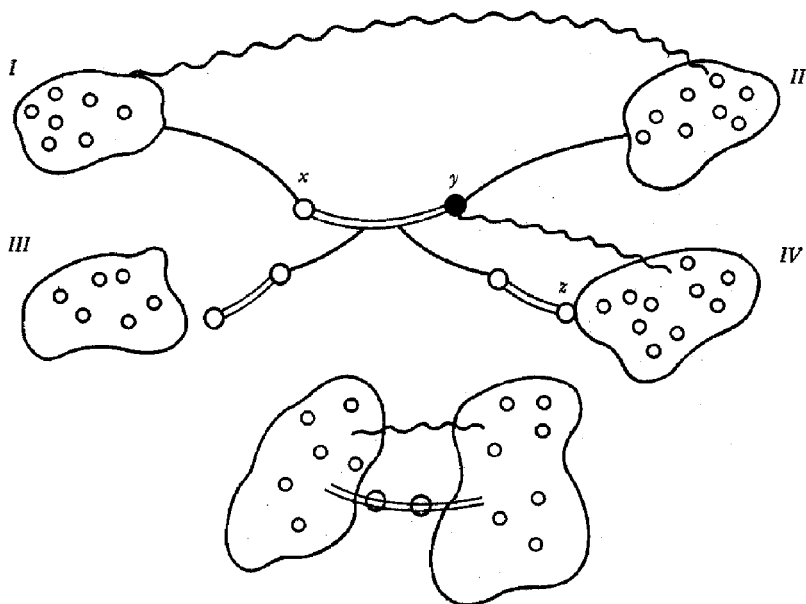
6.3 Diagramming kinship

Anthropology's obsession with kinship has seen it follow many avenues of investigation through decades of research, and the dimensions of interest are far from exhausted (cf., for example, Parkin 1997, Godelier *et al.* 1998, Bloch and Sperber 2002). One intriguing domain which has received little scientific attention is the practice of visually representing kinship not only by anthropologists but also by lay people. As Bouquet (1996: 43) writes, the theoretical

status of the genealogical diagram is ‘rarely considered’. Can a genealogical diagram tell us about the conceptualization of kinship in the imagination of the person who produces it? We occasionally see in anthropological work a non-standard representation offered by the ethnographer (e.g. Rosaldo 1980: 12) or, less often, a representation produced by the people whose kinship system is under investigation. Conklin’s (1969 [1964]: 97) ethnogenealogical method is a ‘highly instructive’ technique of ‘examin[ing] the explicit models constructed by the informants themselves’. Some of the kinship models Conklin elicited from Hanunóo-speakers have visuospatial features. For instance, the Hanunóo talk about degrees of proximity (a spatial metaphor) of interconnectedness of kin with reference to *kāway* ‘flowering branchlets of certain plants’: ‘The Hanunóo refer to relations between such adjacent plant structures when discussing the boundaries of their maximal kin groupings’ and when reckoning permissible marriages (Conklin 1969 [1964]: 101). This in itself is perhaps no different from English-speakers talking of ‘branches’ of a family or organization. More interesting in the present context is a diagram which Hanunóo men spontaneously produced during Conklin’s questioning (Figure 6.1).

Conklin says that drawings such as this did much to clarify his thinking about Hanunóo principles for reckoning kinship (cf. also Deacon and Wedgewood 1934, DeMarrais *et al.* 1992, Rio 2005, Green 2008, *inter alia*). Similarly, Munn (1986 [1973]) reports that during fieldwork on Walbiri iconography she made drawing materials available to Walbiri consultants, eliciting ‘visual texts’ for her research. She says of these representations: ‘I gained not only information on Walbiri iconography, but also a different perspective on the way Walbiri men conceive of their myth and cosmology than I gained from verbal texts’ (Munn 1986 [1973]: xviii). These observations of the special effects of visual sources of information are consistent with the known cognitive effects of diagrammatic representations. Such benefits are also to be gained from analysing hand gestures and other spontaneous bodily movements.

Familiar traditions of kinship diagramming (e.g. standard genealogical trees) furnish enduring external representations designed for display and for physical/spatial and temporal portability across communicative contexts. Like iconographic and semasiographic representations (Gelb 1952, Salomon 2001), they are part of culture’s ‘external symbolic store’ (Donald 1991). Co-speech gesture is, by contrast, an evanescent and apparently less conscious means of diagramming kinship. It does not add to a historically enduring external store but nevertheless does yield a cultural product which forms a focus of joint attention in transitory congregations of interactants (Kendon 1990, Goodwin 2000a), functioning as a kind of external working memory (cf. Baddeley 1986, Donald 1991, Goldin-Meadow 1999, Roth 2000, Roth and Lawless 2000a, b, Clark 2002, Emmorey 2002, Emmorey and Casey 2002).



Kinship is a complex structured conceptual domain that is relatively abstract in spatial content. It has, however, at least two concrete spatial properties, namely, (1) an analogic correspondence between relative height and relative age among kin in childhood and (2) the tendency for greater spatial proximity of certain kin in daily life. Proximity is one-dimensional and therefore spatially somewhat abstract. Furthermore, it is neither a necessary nor inherent property of kinship. Therefore there would seem to be few intrinsic constraints on the way kinship is to be diagrammed. This makes it a useful domain in which to examine the relationship between abstract cognitive structures and concrete spatial/external representations, since the chosen mode of spatial representation will not be pre-empted by direct visual/iconic principles.

Another reason the kinship domain is valuable for the current investigation is that in contrast to other complex conceptual domains, extensive knowledge of kinship systems is not restricted to certain expert individuals in a community. Many facts of kin terminology and kinship-related pre- and proscriptions are common ground in an adult community and are central to social organization and daily life. Much of kinship constitutes lay knowledge, fundamental cultural common ground assumed and accessed in all social interactions.

In order to represent kinship structure visually, it has to be mapped onto spatial structure. While there are few, if any, stable, direct correspondences of kinship relations onto two- or three-dimensional space, the currently prevailing metaphorical principles of spatial mapping seem well motivated. In a typical family tree, greater height corresponds to greater age (an appropriate mapping at least with reference to non-adults). But not all age differences between individuals are made explicit in standard kinship diagrams. Siblings, for example, are often set out at a single level along the lateral axis.

Below I examine some ways in which four semi-urban speakers of Lao externalize their conceptualization of kinship and kinship terminology in informal explanations. Before presenting the data I will establish the background to Lao kinship and kin terminology.

6.4 Lao kinship and kin terminology

Lao kin terms are used as terms of address, as pronouns, and as descriptive terms (Enfield 2007a, 2007b).² Linguistic items used in these three general functions are drawn from a single set, and the selection of a specific term is based not just on the identity of the referent but on the type of referential function being performed. As terms of address, kinship terms are used either alone or as a prefix to the referent's name (Enfield 2007a, 2007b). Thus, if my mother's older sister's name is Seng, I could call her *paa4* 'elder aunt' or *paa4 sèng3* 'elder aunt Seng'. As pronouns, kin terms are used with reference to people higher than oneself.

- (1) *paa4 siø pha2 khòðj5 paj3 talaat5 bòð3*
 peZ IRR take 1SG.P go market QPLR
 'Will aunt ('you'/'she') take me to market?'

² On kinship and kin terminology in closely related systems of Thai speakers, see Keyes (1975), Potter (1976), and Kemp (1984). Writing on Thai, Haas (1978 [1969]: 40) recognizes three ways in which kinship terms may be used: as nouns, as pronouns, and as titles preceding a given name. These are distinctions of grammatical function which crosscut the distinctions focused on here. Kinship notation follows Parkin (1997: 9): F (father), M (mother), B (brother), z (sister), S (son), D (daughter), H (husband), W (wife), P (parent), G (sibling), E (spouse), C (child), e (elder), y (younger). In the kin diagrams, older siblings are placed higher.

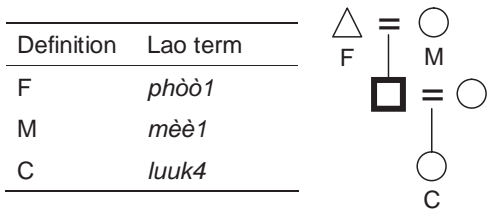


Figure 6.2. Immediate filial relations F, M, and C.

The individual referred to by the kin term *paa4* ‘aunt’ in this example could be the addressee or a third person, depending on context. It is necessary to distinguish between ‘informal’ and ‘official’ descriptive uses of kin terms. A Lao-speaking man may informally describe his father’s older brother’s son as his *qaaj4* ‘older brother’ (Hawaiian style), but he will acknowledge that the proper or official description is *luung2* ‘parent’s elder brother’ (Crow-Omaha style) because of a skewing rule which renders cousin and elder uncle in this case terminologically equivalent (see below).³ In this chapter I focus on the official descriptive use of Lao kin terms – the terms speakers use when they are explicitly and formally defining a kinship relationship (as in ‘That person is my [brother, father, cousin, etc.]’). The basic kinship terminology begins with three terms for immediate filial relations (Figure 6.2). The term *luuk4* ‘child’ is not specified for sex. Specification of sex for this term (and some others) is by suffixation of *saaj2* ‘male’ or *saaw3* ‘female’ (as in *luuk4-saaj2* ‘son’ and *luuk4-saaw3* ‘daughter’).

There are three terms for siblings, each specifying age relative to speaker, with two older-sibling terms distinguished by sex (Figure 6.3). It is impossible to refer to a sibling without explicitly stating whether he or she is older or younger than oneself. The dotted line in Figure 6.3 is of fundamental importance in the descriptive kin terminology of Lao and indeed has a range of consequences in the social life of Lao speakers. It represents the notion of *lun1*, often (inaccurately) translated as ‘generation’. If two people are of the same *lun1*, it means that they are of the very same age – i.e. born in the same year. In English, one is of the same generation as one’s siblings, while in Lao, one’s older sibling is in a *lun1* above and one’s younger sibling is in a *lun1* below. The dotted line divides *lun1 qaaj4/quaj4* ‘older sibling *lun1*’ from *lun1 nò0ng4* ‘younger sibling *lun1*’.

³ The issue of informal versus official uses of kin terms among Lao speakers is worthy of closer attention. Practices of kinship description, as well as members’ knowledge of such practices, are rapidly changing in Lao society and also betray class and regional differences. I hope that an investigation of the broader dynamics of the discourse of Lao kinship will be undertaken elsewhere.

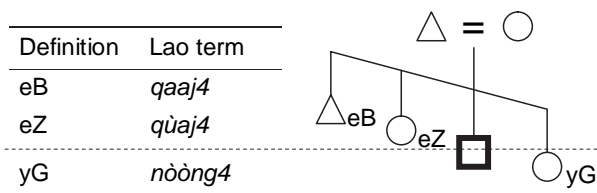


Figure 6.3 Siblings eB, eZ, yG.

Relative age status has a number of consequences for linguistic and other behaviour of siblings toward each other (cf. Enfield 2007a and 2007b: 77ff. for detailed discussion of the information summarized in this paragraph). When referring to an older sibling one typically uses his or her name with the appropriate kin term as a prefix – for example, *qaaj4-phêt1* ‘Phet (older brother)’, *qùaj4-kham2* ‘Kham (older sister)’. When referring to a younger sibling one typically uses his or her name with one of two sex-specific non-respect prefixes, *bak2*- (male) and *qii1*- (female) – for example, *bak2-phêt1* ‘Phet (younger brother or other familiar ‘lower’ male)’, *qii1-kham2* ‘Kham (younger sister or other familiar ‘lower’ female)’. These terms are classificatory in informal address and reference: collateral kin (and also familiar non-kin) are referred to using these terms, with strict reference to status as older or younger than the speaker. Another linguistic practice which makes reference to age differences of siblings is the selection from among a large number of socially marked pronouns (Enfield 2007b: 77ff.). The bare pronouns *kuu3* ‘I’ and *mùng2* ‘you’ are maximally informal and are typically used reciprocally between age-mates who have been friends since childhood or youth (cf. Cooke 1968 on related terms in Thai). They are also common as reciprocal terms of reference between adjacent siblings. They are used **non**-reciprocally by older siblings when speaking to younger siblings. Younger siblings do not use them in return when addressing non-adjacent older siblings unless they intend serious offence. A younger sibling will typically use polite pronouns *khòaj5* ‘I’ and *caw4* ‘you’ when talking to an older sibling. Use of the bare pronouns *kuu3* ‘I’ and *mung2* ‘you’ also extends to interactions with cousins and familiar same-*lun1* social associates as long as the relationship is forged in childhood or youth and has remained continuous since then. As these observations illustrate, the line between younger and older *lun1* is of significant consequence for social and linguistic practice among Lao speakers.

We now turn to descriptive kin terms which denote collateral relations, namely, parents’ siblings and their offspring (cousins). Figure 6.4 shows the six terms for siblings of one’s parents. Each is specified for sex and for age

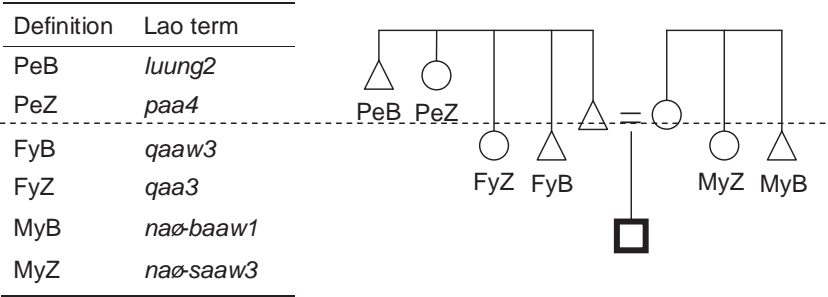


Figure 6.4 Parents' siblings PeB, PeZ, FyB, FyZ, MyB, MyZ.

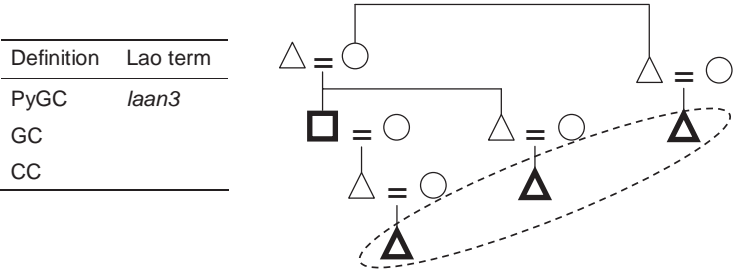


Figure 6.5 The extension of *laan3* (in its official use) is skewed, spanning three generations.

relative to ego's parent, with an added distinction among parents' younger siblings (specifying whether they are on the mother's or father's side). The dotted line separates them into higher or lower *lun1* with respect to ego's parents. One is described in return by all of one's parent's siblings with the term *laan3* (for which sex may be specified by suffixation of *saaj2* 'male' or *saaw3* 'female', as described above for *luuk4* 'child'). The term *laan3*, however, does not simply refer to one's 'nibling'. In its official descriptive use it shows a skewing of three generations, referring to any grandchild (G^{-2}), any child of any of one's own siblings (G^{-1}), or any child of any of one's parents' younger siblings (G^0) (Figure 6.5).

Skewing in the system also occurs in certain extended descriptive uses of the aunt/uncle terms described in Figure 6.4. Official descriptive kin terminology used between cousins is the same as that used between niece/nephews and aunt/uncles (Figure 6.6). If you describe your father's older brother as *luung2* 'older uncle', then officially you describe his son as *luung2* 'older uncle' as well. While your father's older brother's son is in the same generation as you (first cousin in English), he is on a different side of the

Definition	Lao term	Extended reference, via parents' older sibling
PeB	<i>luung2</i>	PeGS
PeZ	<i>paa4</i>	PeGD

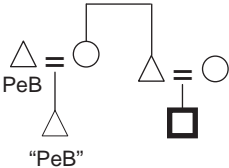


Figure 6.6 Extended reference of the term *luung2* ‘PeB’ for PeGS.

Definition	Lao term	Extended reference, via spouse = parent rule
PeB	<i>luung2</i>	EeB
PeZ	<i>paa4</i>	EeZ
<hr/>		
FyB	<i>qaaw3</i>	HyB
FyZ	<i>qaa3</i>	HyZ
MyB	<i>naø-baaw1</i>	WyB
MyZ	<i>naø-saaw3</i>	WyZ

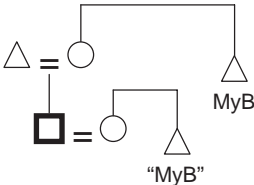


Figure 6.7 Extension of ‘aunt/uncle’ terms to spouse’s siblings. Tree shows example of extended reference of the term *naø-baaw1* ‘MyB’ for WyB.

dotted line, as calculated by the relevant sibling relationship (i.e. the one separating his father and yours). In return, both your father’s older brother and your father’s older brother’s son officially describe you as *laan3* ‘nephew’.

A further type of skewing arises from descriptive uses of aunt/uncle terms (collateral, G^{+1}) with reference to spouse’s siblings. One describes one’s spouse’s siblings using the same terms one’s own child would use, that is, by terminologically identifying one’s spouse with one’s opposite-sex parent (Figure 6.7).

While use of aunt/uncle terms for cousins ‘above the line’ entails being called *laan3* ‘nibling’ by them in return, the descriptive relationship between one’s spouse and one’s sibling is not precisely analogous. For example, a woman officially describes her husband’s older sister as *paa4* ‘older aunt’, while the husband’s older sister describes the woman as *qaa3* ‘younger maternal aunt’ (and not as *laan3* ‘niece’) in return.

To summarize: there is a horizontal line separating ego from ego’s older siblings and younger siblings, respectively. This is observed in the semantics of sibling terms with obligatory specification of age relative to speaker. The three terms are *qaaj4* ‘elder brother’, *qùaj4* ‘elder sister’, and *nòng4* ‘younger sibling’. Terms for parents’ siblings also obligatorily mark age of

the referent individual relative to ego's parents. Cousins are informally referred to by the same terms used for siblings, but officially one has an aunt/uncle-nibling relationship with one's cousins. I describe my parents' older siblings in the same way I describe my parents' older siblings' **children** ('older uncle', 'older aunt'), and they describe me as *laan3* ('grandchild', 'nibling') in return. There is skewing of reference via siblings' marriage as well. One's siblings' spouses are officially described by the same terms as one's **parents'** siblings' spouses. Thus, my older sister's husband is my 'older uncle', while my younger brother's wife is my 'younger aunt'. (Informally, however, my wife's younger sister is simply *nòong4 mia2* 'younger sibling [of] wife'.) A sibling exchange which 'crosses' this line creates a terminological conflict. When a man A marries a woman B, B describes A's older sister C as 'older aunt', and A describes B's younger brother D as 'younger uncle'. If C and D were to marry, the descriptive terminology would clash. One of the two would have to cross the line in order for the terminology to be put right (see below for explication of this point).

As we shall see below, in discussing these and other issues, Lao speakers make repeated reference to the status of individuals as high/big or low/small in the system and use a metaphor of crossing the line between one *lun1* and another.

6.5 Interviews on kinship

In the examples below, consultants were asked to give the meanings and uses of a number of kin terms and to comment on a number of marriage rules and preferences including sibling exchange and second-cousin marriage. The speakers consulted were residents of the city of Vientiane, but all originated in more rural areas of Laos. They were not asked to gesture (gesture was not mentioned), and therefore the gestures they used in response constitute spontaneous and to some degree unreflective spatial mapping of kinship relations.

6.5.1 Example 1: first cousin

Consider one instantiation of the first-cousin relationship, father's elder brother's daughter, illustrated in Figure 6.8.

A speaker explains that she officially categorizes her first cousin as *paa4* 'father's elder sister', while ego is officially categorized as *qaa3* 'father's younger sister' in return.⁴ This is illustrated in Figure 6.9.

⁴ Speakers are not in complete agreement here. Some say that the official description of FyBD is *laan3* 'nibling'.

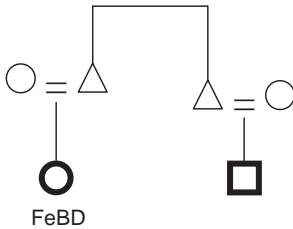


Figure 6.8 Father's elder brother's daughter.

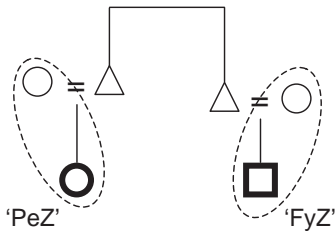


Figure 6.9 A terminological equivalence of cousins and their parents.

Here we see a collapsing of the two generations such that the direct descendent of a parent's sibling is described using the same term as would be used for the parent's sibling, a generation above. The speaker's explanation steps through a full chain of relations in linear fashion. She proceeds along direct filial and sibling relations, first from herself to her father, second from her father to his brother, and third from her father's brother to her father's brother's son (Figures 6.10a–c).

- (2) *khòj5 niø ... pên3 [luuk4 phòd1] –*
 1SG.P TPC COP child father
 'I'm the [child (of my) father]_{Fig. 6.10a –}'
- (3) *laø phòd1 phen1 mii2 [qaaj4] –*
 PRF father 3.P have eB
 'and (my) father, he has an [older brother]_{Fig. 6.10b –}'
- (4) *phòd1 hanø mii2 qaaj4 lèkaø mii2 [luuk4] qiik5*
 father TPC.DIST have eB C.LNK have child more
 '(my) father has an older brother and (he) has a [child] too_{Fig. 6.10c.}'



Figure 6.10 (a) 'I'm the [child (of my) father]-'; (b) 'and (my) father, he has an [older brother]-'; (c) '(my) father has an older brother and (he) has a [child] too'.



Figure 6.11 Diagram of speaker's relation with her father's elder brother's daughter.

The speaker has constructed a mid-air diagram of the form of Figure 6.11. This spatially structured virtual diagram remains available as a target for pointing gestures in the next part of the speaker's explanation. She says, 'We have to refer to his child as ['older aunt']_{POINTS TO FeBD}, and she has to call us ['younger aunt']_{POINTS TO SELF}.'

This segment illustrates a number of features of the spontaneous diagrammatic use of co-speech gesture for depicting kinship. The speaker frames her explanation in terms of four individuals: herself, her father, her father's brother, and her cousin (the target referent). Her purpose in mentioning her father and uncle is to reach the cousin in a linear stepwise series of relationships. She uses pointing gestures to assign the individual referents to

distinct chunks of space, and these spatially anchored reference points remain in place as if drawn on a board. Unlike the speech produced with it, the gesture diagram is spatially multidimensional. The representational nodes are established in linear fashion, but once in place each maintains an intrinsic and direct locational relation to every other node. Because they are not literally drawn or inscribed, to persist as they do they must be mentally represented by the participants in a stable way. As long as the explanation at hand remains the focus of attention in the discourse, the diagram does not evaporate. Eventually, when it is no longer needed, it is erased in a progression to something new. But, while not permanent like a printed diagram, this temporary persistence makes the gestured diagram more than evanescent.

The diagrammatic properties displayed in this example reveal aspects of structure in one fragment of the kinship network which are not available in the linguistic commentary given in (2)–(4).⁵ The two siblings in the structure are both depicted spatially as higher than the speaker while being equal in height to each other. They are in addition laterally distinguished, one on the left, one on the right. The father–daughter relation between father’s brother and father’s brother’s daughter is represented as directly vertical, with daughter below.

6.5.2 Example 2: explaining the term *laan3*

The term *laan3* provides a representational challenge for a speaker trying to explain its meaning in the abstract. As described above (Figure 6.5), the category of *laan3* merges not only lineal and collateral kin but also individuals of three different generations (G^0 , G^{-1} , G^{-2}). The extensional range of *laan3* in a traditional scientific kinship diagram is a diagonal grouping, extending laterally to capture both lineal and collateral relations, and vertically to capture multiple generations. How are such skewed categories conceptualized by speakers? In the examples of diagrammatic representation discussed here, speakers do not represent these categories as diagrammatically singular. Instead, when calculating series of relations, these speakers follow a linear stepwise path of kin relations as in the previous example and enumerate the set of possible referents of a term.⁶ Figure 6.12 shows a fragment of a kinship network, with arrows leading from an individual to those whom he may refer to as *laan3*.

⁵ By this I do not mean that these aspects of structure cannot be expressed in the language at all. I mean that in this case they are not expressed in speech and are retrievable only from the information expressed by hand movements.

⁶ This could be taken to mean that terms with skewed reference such as *laan3* are not thought of by speakers as denoting a unitary extension and that the class is instead thought of as defined by the set of different types of individuals in it. At the same time, it may be the nature of the discourse genre reported on here that elicits this kind of representation.

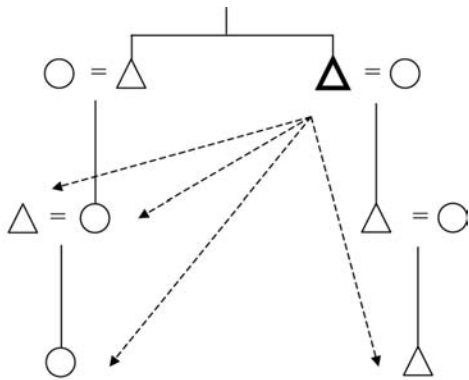


Figure 6.12 Partial extension of the Lao term *laan3*, indicated by arrows.

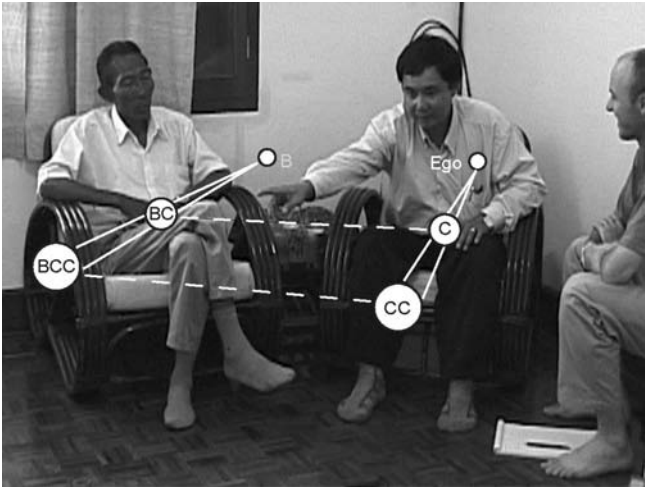


Figure 6.13 Diagram in place for pointing in exposition.

In Chapter 7, below, I describe in detail how one speaker (shown in Figure 6.13) explicates the information depicted in Figure 6.12 using composite utterances of hand gestures and speech. To make reference to the set of relationships depicted by dotted-line arrows in Figure 6.12 the speaker first sets up a kind of diagram with multiple referential nodes at which he can direct eye gaze and finger-pointing gestures. The first sequence in this video-recorded segment involves the use of pointing gestures to set up a virtual diagram which is as if visible and suspended in mid-air (Figure 6.13).

Consider the semiotics of how the non-spatial relations of abstract kinship are mapped in this example onto spatial relations in a publicly shared

diagrammatic sketch pad. The gesture diagram depicted graphically in Figure 6.13 has a number of spatial properties. First, descent leading from the speaker to his child and to his child's child is mapped onto a straight line beginning at the speaker's body and proceeding outwards on a sagittal (away-going) axis (although shifted across to the speaker's left in the gesture space projected by his body). As in the previous example, filiation is mapped onto a non-lateral axis, although this time sagittal rather than vertical. Two lines of descent are presented side-by-side as parallel lines running forward on a sagittal plane, symmetrically occupying left and right halves of the speaker's projected gesture space.

As in Example 1, this symmetrical use of left and right space is appropriate for representation of two comparable but contrasting entities. It is not that the two must be represented in this way, but for purposes of what is being said about them at this point in the discourse they are being lined up side-by-side. Here, two brothers are placed symmetrically on a lateral axis (just like the father and uncle pair in the previous example), and the collateral lines of descent are correspondingly symmetrical. Thus, the spatial properties of this spontaneous representation directly preserve logical properties of the set of relationships being described. These spatially represented properties are not represented in the accompanying verbal message and are not available in this case without access to the manual/visuospatial information. Here we see hand gestures doing for us what Conklin's ethnogenealogical diagram (Figure 6.1) may have done for him. Once this diagram is set up, it has some degree of persistence despite its lacking real visibility. The speaker uses the diagram like a drawing on a board, employing pointing gestures to link it with the verbal exposition.

In the following example, the speaker states that he describes his brother's son as *laan3* (the brother's name is Naak).

- (5) *khòj5 hòng4 [luuk4] qajø-naak4 kaø ñang2 pên3 laan3*
 1SG.P call child EB-N. T.LNK still COP laan3
juu2
 FAC.WEAK
 'I still call Naak's [children] *laan3*, nevertheless.'

Next, the speaker states that his own child is described by his brother as *laan3* also.

- (6) *qajø-naak4 hòng4 [luuk4 khòj5] kaø pên3 laan3*
 EB-N. call child 1SG.P T.LNK COP laan3
khùj2 kan3
 like COLL
 '(and) Naak calls [my children] *laan3*, too.'

Despite being literally invisible, the diagrammatic structure established in this sequence (depicted graphically in Figure 6.13) functions like a real diagram, an artefact in the perceptual common ground. It enables sustained reference to multiple individuals during a stretch of talk, functioning as external working memory. The structure provides targets for pointing gestures which pick out individuals in the network. As noted above, the term *laan3* (child's child, sibling's child, parent's younger sibling's child) poses significant referential complexities, requiring the speaker to link two individuals by five relationships with four intervening individuals (Figure 6.13). With the representational aid of hand movements, eye gaze, and other mechanisms of nonverbal behaviour, this speaker manages well.

6.5.3 Example 3: marriage between first cousins versus between second cousins

The next segment continues directly from the previous one, with the speaker moving on to a new sub-topic in the discourse. Once the diagram in Figure 6.13 is established in collectively accessible space and, indeed, in the collective discourse record, the speaker is able to exploit it in making further comments about kinship relations between the individuals depicted. He now wants to say that second cousins may marry as long as the male of the pair is descended from the older of the two sibling ancestors. The male second cousin should be from the higher side if the two are to marry (regardless of the relative age of the pair to marry). The parents will be first cousins, such that the girl's parent classifies the boy's parent as *luung2* 'older uncle' or *paa4* 'older aunt' (because of the girl's parent's parent's being younger than the boy's parent's parent). In the following example, the speaker states that his own grandchild can marry his brother's grandchild (the two being second cousins), since they are sufficiently 'far apart'. He points to the distant nodes which stand for the second-cousin pair (BCC and CC in Figure 6.13), saying, '(They) get married, those ones – being yonder, they can marry, since it's far.' Next, the speaker refers to the 'closer' pair, that is, the first-cousin pair, his own child and the child of his brother. As he does so, he draws his arms in closer to himself and gestures as if pinching the two nodes for the first generation below (BC and C in Figure 6.13), saying, 'If they're close together, like so, they're not allowed to marry' (see Chapter 7, below, for detailed explication of this example).

This example shows gesture-generated diagrams being used as cognitive artefacts. Interpretation of the symbolic indexical expressions 'those ones' and 'like so' would be impossible without access to the spatial diagram and the referential value of each of its nodes. The speaker's composite utterance design accords critical responsibility to the virtual diagram for correct

interpretation of the linguistic component. As the speaker refers verbally to ‘those ones’, he points to pieces of space which, while having been accorded reference earlier, were, just prior to the relevant moment, literally empty. The speaker makes no check for comprehension of this reference and shows no evidence that he expects any problem in understanding. He is treating the (in fact invisible) diagram as an unproblematic source of information to his addressees. This indicates that he assumes that his addressees are cognitively maintaining the diagram through time, across segments in which there is no pointer to the space, no physical realization of a referent in that space, and no referring speech. The same is true in Example 2, above; the speaker points to something which at the preceding moment had no corporeal form, no pointer oriented to it, and no verbal reference. That this passes without comment indicates that the virtual diagram, as a cognitive artefact required for the representational and interpretive task at hand, is assumed by participants to be collectively entertained.

6.5.4 Example 4: one descending generation on a vertical line

In the data set from which the present examples are drawn, Lao speakers are never observed to map a relation of filiation onto the lateral axis when using gesture for spatial representation. In Examples 2 and 3, above, the speaker uses the sagittal axis as a line of descent, mapping the parent–child–grand-child line onto a vector going forward from the centre of the body. In the following segment, filiation is similarly mapped onto a non-lateral axis but this time on a vertical line (cf. Example 1, above).

The speaker begins an explanation of sibling terms. He will eventually concentrate on the kin terms used between four brothers, but first he establishes the relation of the father to the first son, the speaker’s own older brother (see Chapter 7 for detailed description of the sequence). We begin here at a point where the speaker has introduced his own father as a discourse topic. Figure 6.14a shows the speaker representing his father (as he mentions him in speech) by finger-pointing to a space at belly height, directly in front of his body. Having introduced his father into the discourse and established a point in space to refer to him, the speaker then refers to the father’s first son (i.e. the speaker’s own older brother), pointing to a space vertically beneath the point on which he had mapped the father. He says, ‘This one here (i.e. the father [cf. Figure 6.14a]) – is the one who brings about the creation – is the one who brings about the creation (of) the- this [*qaaj4*] (the oldest brother)’ (cf. Example (14) in Chapter 7).

It would not be possible to resolve reference of the symbolic indexical ‘this one here’ (*phuu5 nii4 niø*) without reference to a conceptual representation of the gesture-generated virtual diagram. That neither speaker nor addressee



Figure 6.14 (a) Speaker points (and gazes) as he refers to ‘father’; (b) Speaker points as he refers to ‘first son of father’; (c) Relation of filiation represented on a vertical line in gesture space.

shows any evidence of a real or expected problem indicates that the diagram is useful for both utterance production and comprehension and the joint attention that unite these. Figures 6.14a and 6.14b show the relation of filiation represented on a vertical line, depicted graphically as in Figure 6.14c. And again, certain information about the structural properties of kinship relations is depicted in the spatial construction of the hand-gesture sequence and not in the accompanying talk (e.g. the relation of vertical placement of father over son is not encoded in speech; see below for further discussion).

6.5.5 Example 5: sibling relations: four brothers on a diagonal line

Lao kinship terminology represents kinship relations as inherently hierarchical, based on sibling birth order. Unsurprisingly, Lao speakers in these interviews employ differences in height in spatially representing sibling relations. Figure 6.15 depicts a virtual diagram illustrating four brothers, of whom the speaker is second oldest (see Chapter 7, below). The speaker uses nodes of the diagram as targets for pointing gestures while referring in the speech to the individuals who are mapped onto those locations.

As in the previous example, these pointing gestures are distributed over time, with nodes on the virtual diagram being temporarily left unattended and the speaker apparently assuming no problem in their interpretation. The speaker differentiates visually between siblings by combining lateral and vertical dimensions in space. The result is a diagonal line. The vertical dimension represents relative age (i.e. the older individual is higher), and the spacing of siblings along the lateral dimension represents their side-by-side status as heads of new collateral lines of descent (see discussion of Examples 1–3, above). The conceptual relations manifest as a diagonal axis are directly represented in the spatial structure of the hand-gesture diagram and are not directly available in the coded linguistic structure supplied.

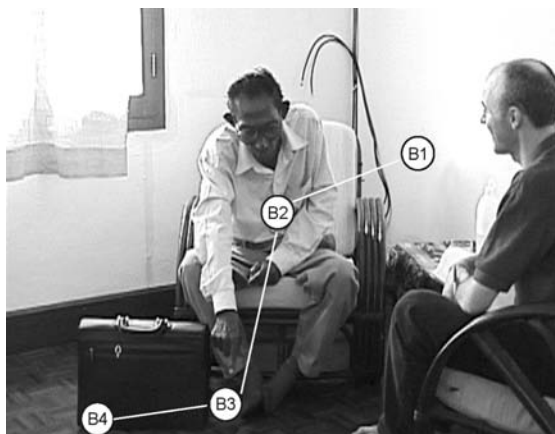


Figure 6.15 Relation between four brothers represented on a diagonal line in gesture space.

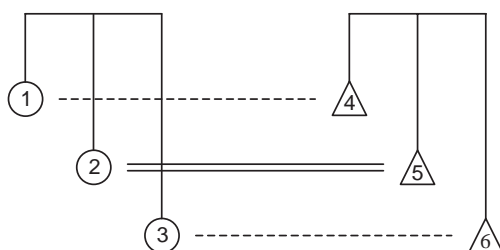


Figure 6.16 Some possible sibling exchanges among Lao speakers (dotted lines); it is permitted for two older or two younger siblings of a married couple to wed. (Double lines indicate that 2 and 5 are married.)

6.5.6 Example 6: sibling exchange and relative height, I

A constraint on sibling exchange described by Lao speakers can be explained with reference to the diagram in Figure 6.16.

A woman's sibling may marry her husband's sibling only if they are both older siblings or both younger siblings of the woman and her husband, respectively (regardless of the relative ages of the pair to wed). In Figure 6.16, the permitted marriages are $1 = 4$ and $3 = 6$. In the following excerpt, the speaker uses hand movements to accompany his spoken explanation of the principle depicted. He says that older siblings of a husband and wife can

marry, since they are both on the same side of (i.e. above versus below) the line defined by sibling birth order:

- (7) [phen1 pên3 qàaj4 – pên3 qaaj4 khòòng3 phuak4 haw2
 3SG.P COP eZ COP eB of group 1.F
 laø mèèn1 qaw3 daj4]
 PRF COP marry CAN
 ‘[They are older sister – (and) are older brother of us, and so they can get married.]Fig. 6.17a’

Figure 6.17a (below) shows the speaker using an accompanying gesture in which height is employed to represent the status of the two siblings as above the level of the speaker and his wife (i.e. as older siblings or equivalent). Further, the equivalence in height of the two hands iconically portrays the equivalent status of the two siblings (as both at the same level). These notions can be taken as conveying information about Lao kinship structure but not encoded in the accompanying speech.

6.5.7 Example 7: sibling exchange and relative height, II

Soon after the segment described in Example 6, the same speaker elaborates on possible sibling-exchange marriages, stating the two logical possibilities – two older siblings marrying and two younger siblings marrying:

- (8) khan2 vaa1 [ñaj1] kaø ñaj1 bet2-
 if COMP big T.LNK big all
 khan2 vaa1 [nòj4] kaø nòj4 bet2
 if COMP small T.LNK small all
 ‘If (one is) big, Fig. 6.17b (they should) both (be) big, if (one is) small, Fig. 6.17c (they should) both (be) small.’



Figure 6.17 (a) ‘They are older sister- (and) are older brother of us, and so they can get married’; (b) ‘If (one is) big (they should) both (be) big’; (c) ‘If (one is) small, (they should) both (be) small.’

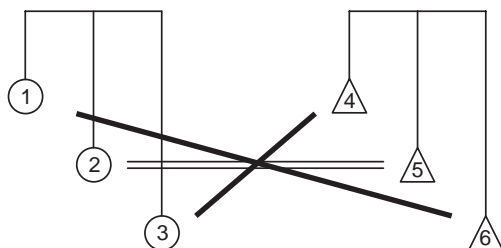


Figure 6.18 Constraints on sibling exchange among Lao speakers; it is not permitted for an older and a younger sibling of a married couple to wed. The proscribed marriages are indicated by bold lines.

Again, the speaker uses equivalence in height of the two hands to represent equivalence in height of the two suitors and also directly represents spatially the status of the pair as above or below the speaker himself (taking his upper chest as a reference centrepont). In addition, in this sequence the lateral dimension is also exploited. The high gesture in Figure 6.17b is to the speaker's left; the low gesture in Figure 6.17c is to the speaker's right. This use of laterality is a typical way of representing a contrast or comparison of two situations (see above). The information conveyed by this placement of referents on the left and right of gesture space is not provided in the concurrent speech.

6.5.8 Example 8: sibling exchange and relative height, III

We now turn to the proscribed sibling exchanges, in which the line is crossed (i.e. a person's younger sibling marries the older sibling of the person's spouse [cf. Figure 6.16]). The rules outlined above may be consulted to elucidate the problem here (Figure 6.18). If 3 and 4 marry, 2 has a problem with respect to her relationship with 4. As husband of 2's younger sister, 4 would be 2's *naø-baaw1* 'younger uncle', but as older brother of 2's husband, 4 would be 2's *luung2* 'older uncle' (see Figure 6.7).

The next data excerpt focuses on a specific instance of the type of problem illustrated in Figure 6.18, namely, a man's younger brother marrying the man's wife's older sister (Figure 6.19).

The speaker begins with a reference to his own wife, finger-pointing to the front as he refers to her in the speech (Figure 6.20a). He then refers to her sister, his own sister-in-law, raising his index finger (Figure 6.20b) and locating her at a point higher in space than her sister (his wife).⁷

⁷ This is a clear example of gesture's running well ahead of speech in time of production (McNeill 1992: 25). The pointing gesture referring to the wife's older sister comes well before its lexical affiliate *qàaj4* 'older sister'.

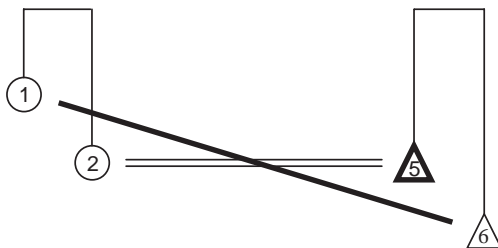


Figure 6.19 Example of non-permitted marriage: younger brother = wife's elder sister. Speaker in Figure 6.20 takes the perspective of 5.



Figure 6.20 (a) 'But – the one who is – my wife . . .'; (b) '... she has an older sister – right?'; (c) 'But [I] have a younger sibling – ...'

- (9) *tèl1 vaa1 [-] phuu5 qanø – mia2 [khòòng3*
 but COMP person HES wife of
khòj5] hanø laø laaw2 mii2 qàaj4 – nòq1
 1SG.P TPC.DIST PCL 3SG.FA have ez QPLR.AGREE
 'But – the one who is – my wife_{Fig. 6.20a}, she has an older sister –
 right?_{Fig. 6.20b}'

He now refers to his own younger brother, whom he locates in space at a point lower than the level of himself and his wife, pointing down close to his side (Figure 6.20c).

- (10) *tèl1 vaa1 [khòj5] phat1 mii2 nòòng4*
 but COMP 1SG.P T.LNK have yG
 'But [I] have a younger sibling –_{Fig. 6.20c}'

Now entertaining the idea of his younger brother marrying his wife's older sister, the speaker once again refers to the wife's older sister. In doing so, he finger-points (Figure 6.21a) to the already established point on his hovering virtual diagram, to the front and higher in space than the point reserved for his wife.



Figure 6.21 (a) ‘... (for him) to marry (her) – is not possible’; (b) ‘If my younger sibling ...’; (c) ‘... would marry the younger sibling of my wife ... (that’s) possible’.

- (11) *siø paj3 [qaw3] hanø naø – bøø daj4*
 IRR go marry TPC.DIST TPC NEG CAN
 ‘(for him) to [marry (her)]_{Fig. 6.21a} – is not possible.’

The speaker refers by finger-pointing to a referent which is not mentioned in the accompanying speech and is not visible. The gesture in Figure 6.21a is directed toward the same point in space as that in Figure 6.20b. There has been a break between these two moments during which no reference, verbal or gestural, was made to the relevant point in space (Figure 6.20c). In Figure 6.21a, the speaker points to the temporarily unattended node without showing any sign of expecting trouble in interpretation, and indeed no such trouble arises. The information in the gesture is not only for the addressee but is being treated by the speaker as **usable by** the addressee.

The speaker elaborates further, now clarifying that if his younger brother were to marry any sister of his wife, it would have to be the younger sister (cf. Figure 6.16). He first points to the younger brother, established in Figure 6.20c and being returned to now (Figure 6.21b). In first mentioning the wife’s younger sister, he makes a new pointing gesture (accompanied by eye gaze [Figure 6.21c]), not just pointing to but **creating** the diagrammatic node corresponding to the wife’s younger sister.⁸

- (12) *khan2 [nòòng4 khòòng3 khòòj5] siø [qaw3 nòòng4]*
 if yG of 1SG.P IRR marry yG
khòòng3 mia2 hanø... daj4
 of wife TPC.DIST CAN
 ‘If my younger sibling_{Fig. 21b} would marry the younger sibling of my wife ... (that’s) possible_{Fig. 6.21c}.’

⁸ Such creating gestures are described as baptismal by Haviland (2000: 20).

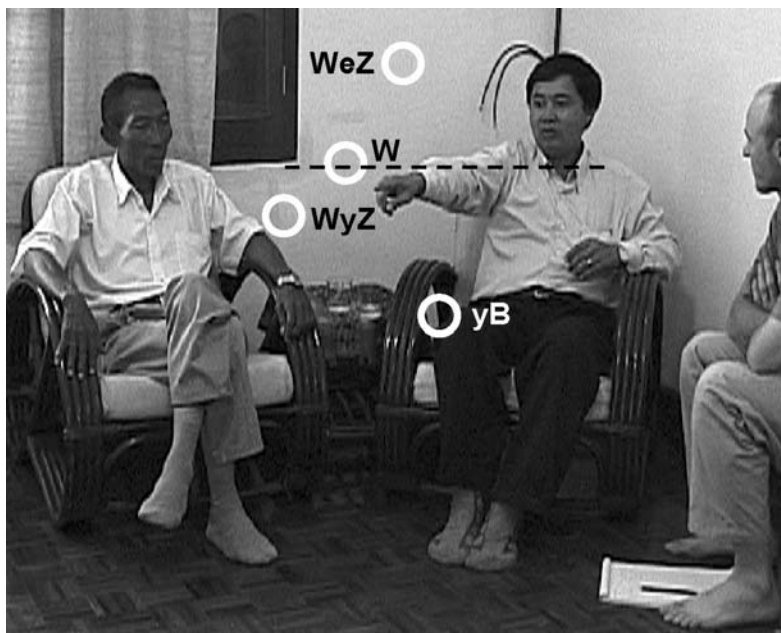


Figure 6.22 Proscribed marriage between wife's elder sister and younger brother would cross 'the line' implicit in the gesture diagram (represented here as a dotted horizontal line). Permissible marriage between wife's younger sister and self's younger brother does not.

The diagram this speaker has now produced can be made explicit as in Figure 6.22.

The semiotic use of space in this sequence is rich, achieved not only by hand movements but also by eye gaze. Height is again used to represent status as to individuals being above or below the line established by sibling birth order. This and other spatial features of the display are not encoded in the speaker's verbal descriptions. The speaker neither expects nor creates problems of interpretation in gesturally referring to referential points in space which have been temporarily left empty.

6.5.9 Example 9: sibling exchange and relative height, IV

In a further example, a different speaker is discussing a situation similar to that in Figure 6.19, depicted as in Figure 6.23.

Here, the speaker (taking the perspective of 5) is talking about the hypothetical possibility of his wife's older brother (1) marrying his younger sister (6). He begins with pronominal reference to his wife's older brother (using

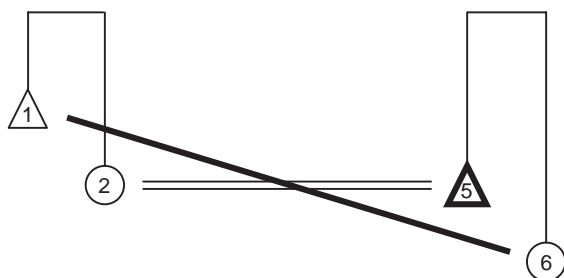


Figure 6.23 Example of non-permitted marriage: younger sister = wife's elder brother. Speaker takes the perspective of 5.

the polite third person singular form *phen1*), accompanying the speech with a pointing gesture high and far to his right (Figure 6.24a).

- (13) *phen1 pēn3 sà4 [suung3] haw2 kheej2 phen1*
 3SG.P COP bloodline high 1.FA EXP 3SG.P
hòòng4 naø-baaw1
 call MYB
 'They are of a high bloodline; we're accustomed to their calling us
 "younger uncle" Fig. 6.24a.'

Then, he refers to his younger sister, tapping on the floor low and far to his left (Figure 6.24b).

- (14) *laø phen1 maa2 qaw3 kap2 phuu5 [tamø-tam1]*
 PRF 3.P come marry with person RDP.A-low
phuak4 haw2 niø
 group 1.FA TPC
 'And then they come and marry "us" low ones here Fig. 6.24b.'

In this sequence we again see height of hand gestures (in relation to the speaker's own torso level) being employed to indicate position of kin as above or below the line established by sibling birth order. In this example, height is encoded in the speech, but the separation of the two related individuals on a lateral axis is not.

6.5.10 Example 10: sibling exchange and relative height, V

A final excerpt on sibling exchange illustrates a striking use of the hands in a composite utterance. After explaining the problem of marriages such as that depicted in Figure 6.23, the speaker wants to articulate further the nature of



Figure 6.24 (a) ‘They are of a high bloodline; we’re accustomed to their calling us “younger uncle”.’ (b) ‘And then they come and marry “us” low ones here.’



Figure 6.25 (a) ‘It’s – it’s crossed’ (b) ‘together–’.

the problem. He uses the word *khuaj3* ‘crossed’ to describe his assessment of the proscribed sibling exchange, while exploring this notion manually, circling one hand around the other, as if searching for the right representation (shown in Figures 6.25a–b).

- (15) *man2- man2 [khuaj3] [kan3] –*
 3.B 3.B crossed COLL
 ‘It’s – it’s crossed_{Fig. 6.25a} together_{Fig. 6.25b} –.’

He concludes as follows, settling on the striking posture depicted in Figure 6.25c, vividly evoking the idea of the illicit couple’s having ‘crossed the line’.

- (16) [*man2 siø*] *pên3 – khuaj3 bèèp5 sii4*
 3.B IRR COP crossed like thus
 ‘It would be – crossed like this_{Fig. 6.25c}.’



Figure 6.25c ‘It would be – crossed like this.’

6.6 Representational and conceptual properties of gesture diagrams

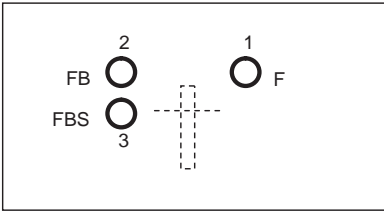
In the absence of telepathy, communicating an idea to another person involves both internal (cognitive) and external (perceptible) processes. In order to make our inner states known, we must produce external representations which are perceptually available to others (Miller 1951: 3). And in order to know the inner states of others, we must perceptually access external representations and transform them into corresponding internal representations (Reddy 1979, Hutchins and Hazlehurst 1995). These processes of conceptualization, articulation, perception, and comprehension are tightly intertwined in the interactional achievements of discourse (Schegloff 1982, Goodwin 2000a). Bodily movements including co-speech hand gestures are resources for solving these conceptual problems of both production and comprehension in communicating complex ideas. External representations feature not just in the range of cognitive and practical tasks (Goody 1977, Larkin and Simon 1987, Button 1990, Norman 1991, Zhang 1997, Heath and Luff 2000, Roth 2000, Roth and Lawless 2002a) but also in the ‘technology of conversation’ (Sacks 1984: 413). In the Lao kinship examples, speakers not only choose to accompany their speech with visual representations but also choose, with some consistency, specific forms of visual representation. What are the implications of these representational choices?

The same piece of information may be externally represented in different forms, and the given form of an external representation can have a significant effect on how it is perceived and, in turn, on how the information it conveys is cognitively processed. A number of taxonomies of types of representation have been proposed (cf., e.g. Kosslyn 1978, Palmer 1978, Larkin and Simon 1987). Examining types of paper-and-pencil external representations, Larkin and Simon (1987) distinguish between sentential and diagrammatic modes. Sentential representations are data structures ‘in which elements appear in a single sequence’ (Larkin and Simon 1987: 68). These sequences are either natural-language sentences or formal translations of them. Diagrammatic representations, by contrast, convey information via data structures ‘in which information is indexed by two-dimensional location’ (ibid.). Pieces of information are given spatial locations relative to each other and can be searched via the ‘computationally easy process’ of change of attention to adjacent locations, that is, by direct visual inspection (Larkin and Simon 1987: 80; cf. Goody 1977: 134, Harris 1986: 139). Spoken and visual representations have significantly different affordances for fundamental problem-solving cognitive operations such as search, recognition, and inference (Larkin and Simon 1987: 69; cf. Glenberg and Langston 1992, Zhang 1997). Visual representations have a number of cognitive advantages.

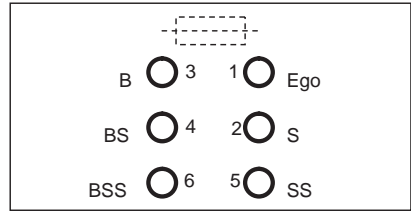
The external memory function of diagrams arising from simultaneous perceptual availability of multiple items of information can be considerably enhanced if the mapping between a form of representation and its value is a natural one. The naturalness of a mapping is related to its directness, which ‘can be measured by the complexity of the relationship between representation and value, measured by the length of the description of that mapping’ (Norman 1991: 28; cf. also Palmer 1978: 270–271). Intrinsic representations (Palmer 1978) are natural ones, allowing representational relations to be ‘efficiently read off from the display’ without the need for them to be ‘inferred from symbolic descriptions and coordinates’ (Oestermeier and Hesse 2000: 93; cf. Norman 1991: 30–31). What types of information are intrinsic to the diagrammatic hand gesture sequences described above? Three representational properties of gesture sequences are of special importance for understanding the nature of hand gestures and other bodily movements in their role as cognitive artefacts: Virtual reality, enchronic linearization, and a bodily origo for three-dimensional spatialization.

Hand-gesture sequences have an extraordinary ability to give rise to virtual structures. The illustrations in Figure 6.26 do not depict the gestures themselves; they represent what the speakers may be conceived to have pointed to in the sequences described.

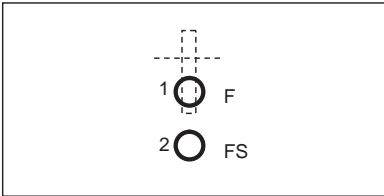
The gestures themselves are analogous not to the weather map but to the stick which indexically connects the weather reporter’s map with his



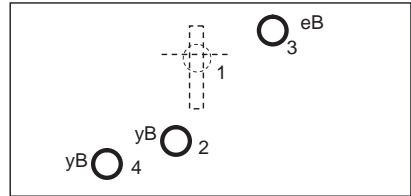
a. Front view of diagram, Example 1



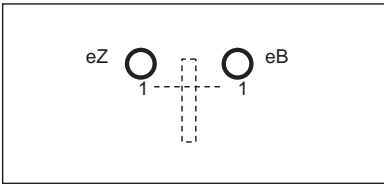
b. Overhead view of diagram, Examples 2/3



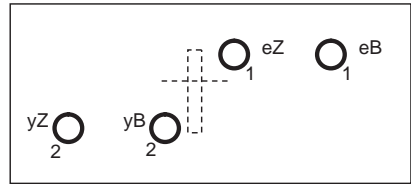
c. Front view of diagram, Example 4



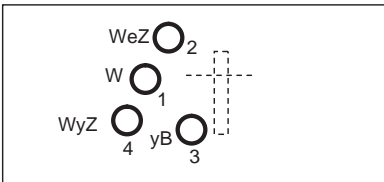
d. Front view of diagram, Example 5



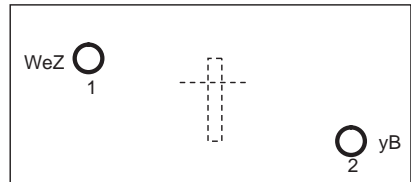
e. Front view of diagram, Example 6



f. Front view of diagram, Example 7



g. Front view of diagram, Example 8



h. Front view of diagram, Example 9

Figure 6.26 Abstract representations derived from hand-gesture diagrams described in Examples 1–9. The dotted-line figure represents the torso of the speaker (with line crossing at upper chest height); circles represent chunks assigned reference to individuals in the discourse. Numerals represent the order of appearance of nodes in the series of gestures.

conventional, symbolic (linguistic) commentary. And, most of the time, the initial setting up of the gesture diagram is performed by baptismally pointing as if the referent were already there (Haviland 2000). A weather map is a persistent physical object, while gesture diagrams are transient and virtual. In the contexts in which they occur they are almost entirely imagined. Nevertheless hand gestures have the properties of cognitive artefacts. That is, they are 'designed to maintain, display, or operate upon information in order to serve a representational function' (Norman 1991: 17). But the diagrams they create are cognitive artefacts of a special type, context-bound and interactionally achieved. The gesture diagrams depicting Lao kinship arise from ordinary speakers' resourceful attempts to articulate abstract and complex sets of relations. With speech as an attentional and symbolic anchor, these speakers introduce visually signifying material into the environment of their interlocutors, a prime field for entry of information into the conceptual common ground (Clark 1996). The result is the maintenance of virtual diagrams in the collaborative and co-present imagination, grounded by passing indications in a common perceptual field.

Zhang (1997) defines external representations as 'the shapes and positions of the symbols, the spatial relations of partial products, etc., which can be perceptually inspected from the environment' (Zhang 1997: 180) Where in the environment is the gesturer's diagram? With the assistance of hand gestures and eye gaze, speakers bestow representational values on transparent chunks of space (see also Chapters 5 and 7; cf. Liddell 1995, 2000, 2003). These collectively imagined chunks maintain their absolute positions, creating an impression of a cohesive spatial structure as if actually present and perceptually available. Like enduring structures such as printed diagrams, gesture diagrams are externally anchored and simultaneously available to more than one (perceiving, thinking) person, thereby enabling distributed representation (Norman 1991, Hutchins 1995). This kind of simultaneity and reciprocity of production and comprehension is the essence of distributed cognition in everyday interaction (Goodwin 2000a).

In the kinship examples, referential values are mapped onto spatial locations without being continuously available for perceptual inspection. On repeated occasions, speakers leave points and later go back to them, yet show no evidence that they expect trouble in interlocutors' understanding due to the momentary lack of physical form accorded to that element in the structure. They apparently presume that the diagram is being maintained in their interlocutors' imagination. While a conceptual mapping between signifiers and signifieds is generously supported by having the relevant signifying material simultaneously available in perception, the gesture examples show that this persistent artefactual material is not even necessary. This is of some

consequence for those who have insisted on the direct perceptual effects of external representations.

A second intrinsic representational property of gesture sequences is the inescapable temporality of their production and comprehension (Goodwin 2002). While drawn or printed diagrams are often accessed at a time and place other than the time and place they are produced, gestures seldom have context-free form. While the information in a drawn or printed diagram can be scanned in any order (despite certain orders suggested by symbols such as arrows) and indeed can be perceived all at once, a sequence of gestures is enchronically interpreted as a series of movements supplied in the order of their production through time. (Accordingly, Figure 6.26 specifies the order of appearance of each node on each diagram.) Gesture, like language, is forced to linearize when it reaches its representational limits (cf. Chapter 5, above).

What is the function of a speaker's choice of relative ordering in the enchronic supply of information? First, as for speech, temporal succession in the delivery of informational chunks affords (indeed forces) asymmetries in the global structuring of information in discourse. What comes first can be assumed to be conceptually active for some period while new pieces of information are added to the discourse record (Chafe 1994). A speaker's decision to express *p* before *q* may be due to his or her foreseeing that *p* is a prerequisite for comprehending *q*. This principle is likely to have applied in a number of the above examples (e.g. Example 1, Figure 6.26a). Alternatively, a speaker's decision to express *p* before *q* may simply be a reflection of his or her having thought of *p* first. As Efron puts it, one function of co-speech gesture is to provide a chart of the course of talk, 'outlining the logical itineraries of the corresponding ideational processes' (1972/1941: 122). Chapter 7, below, presents a detailed description of the sequences summarized above as Examples 2 and 3, detailing speakers' encounters with technical problems in representation arising from not having foreseen the need to incorporate certain features into the diagram at a stage when its general structure had already been established. As Chapter 7 will show, speakers can employ systematic mechanisms for editing gesture diagrams during the course of production.

A second function of choices in ordering the enchronic supply of information in gesture sequences is (again) also observed in speech. The temporal order of expression of events may directly correspond to the temporal order of the occurrence of those events (Haiman 1985). In Figure 6.26c (Example 4), for example, the speaker introduces higher generations before lower, directly representing their order of appearance in the world. The order of introduction of individuals in Figure 6.26a (Example 1), in contrast, appears to reflect the speaker's path of calculation, linearly from A (self) to B (A's father), then to C (B's brother), then to D (C's child).

A third intrinsic representational property of hand-gesture sequences arises from the status of the human body as a visual and proprioceptive centre for anchoring meaningful oppositions. Our expressive signs, whether performed by voice or by arms and hands, are projected from a spatiotemporal origo (Bühler 1982 [1934], Hanks 1990), a deictic zero for the I/here/now, and a grid centre for spatial oppositions including high/low, left/right, and forward/back. This perspective affords an analysis of the semiotic mappings illustrated in Figure 6.26. For example, speakers consistently use height not only to mark relations between different nodes on their diagrams but also to mark relations between those nodes and the speakers themselves. Apart from Figure 6.26b, all the examples feature the use of height to distinguish older individuals from younger individuals. None of the linguistic items which express these kin relations make explicit reference to height (although height is sometimes referred to in accompanying speech; cf. Examples 13 and 14, above). In Figure 6.26b, the relation of relative age is mapped onto distance away from the speaker. When two individuals are of the same age-rank for the purpose of what is being said, this is conspicuously represented by mapping them onto points at identical height. This is illustrated in Figures 6.26a (FB and F), 6.26e (eZ and eB), and 6.26f (the pair eZ and eB and the pair yZ and yB).

Now consider the use of laterality. In Figure 6.26a, the speaker places two siblings – her own father (F) and his brother (FB) – in symmetrically opposed positions on each side of her body. The lateral axis is used in this case to situate two entities which are of the same kind but in some relevant sense different, just as we might make a two-handed laterally symmetrical gesture to express a paradigmatic contrast more generally (Kita *et al.* 2001: 131–134), as if weighing up two comparable but distinct entities. Siblings are repeatedly distinguished by lateral opposition (Figures 6.26a, b, d), a strategy accentuated in Figure 6.26b, by the representation of lines of collateral descent running forward in parallel from their source in two siblings. And when direct filiation is represented, in none of the cases examined here is it depicted using a lateral opposition in space. In Figures 6.26a and 6.26c, the relation of direct descent in kinship is mapped onto a spatially vertical relation. In Figure 6.26b, it is mapped onto the sagittal axis. Perhaps the reason speakers avoid mapping filiation laterally is that, as noted already, a representation anchored in the lateral divide across the centre of the body – an archetypal symmetrical object – would suggest some near-equivalence or comparability between the two referents which might be inappropriate in the context.

These three properties of gesture diagrams – virtual reality, enchronic linearization, and a bodily origo for three-dimensional spatialization – set them apart from garden-variety tools for cognition such as light switches, dashboards, and balance sheets. We now consider more directly their status as cognitive artefacts.

6.7 Gestures and gesture diagrams as cognitive artefacts

Research on cognitive artefacts acknowledges the role of the human body in determining affordances of cultural and technological artefacts – e.g. door handles should be at about the right height to be comfortably reached by a forward-reaching hand when standing (Norman 1988; cf. Wierzbicka 1985). But the idea of the body itself as a cognitive artefact has seldom been made explicit. Research on gesture has established that speakers can use hand movements to spatialize abstract concepts (e.g. McNeill 1992, 2000), to assign reference to seemingly empty space (e.g. McNeill *et al.* 1993), and to express information not encoded in accompanying speech (e.g. Goldin-Meadow 1999, 2003a). But the specific link to research on cognitive artefacts and, in particular, on the special semiotic affordances of gesture diagrams has been relatively unexplored (but see Emmorey 2001, Liddell 2003 on sign language, Roth 2000, Roth and Lawless 2002a, b, on gesture in educational contexts).

A connection between research on gesture and cognitive artefacts is fruitful for a number of reasons. As traditionally described, cognitive artefacts not only assist in solving cognitive tasks, they also change the nature of those tasks (Norman 1991). Visual representations in the task of communication – both for producers and for interpreters – are an important example. Scholars like Goody (1977), Donald (1991), and Olson (1994) suggest that certain types of enduring external representation have historically transformed the way humans think. These developments did not simply involve writing down speech. They took advantage of properties of the written medium which speech itself could never have afforded. But prior to the profound historical developments which writers such as Goody (1977) describe, language had already altered the way we think as a species (Noble and Davidson 1996, Deacon 1997). And further to this, we must be reminded that language is not speech alone (Kendon 1972, 2004, McNeill 1985, 2005). Linguistic utterances routinely take semiotically composite form, combining speech with visible bodily behaviour in holistic utterances. Human gesture is far from the primitive ‘call system’ that some writers suggest (e.g. Burling 1993).

There are three main lines of evidence and arguments in support of the cognitive artefact claim for co-speech bodily actions in the Lao kinship interviews, discussed above: (1) recipient orientation, (2) referential dependency, and (3) semiotic compositeness of moves. Let us consider these in turn.

Speakers’ diagrammatic gesture sequences are studiously and deliberately **presented to recipients**. As Goodwin (2000a) forcefully contends, ‘participants don’t produce talk or build action into the air, but instead actively work to secure the orientation of a hearer . . . and design the current action and utterance in fine detail for the particularities of the current addressee’ (2000a:

1499; cf. Bavelas *et al.* 1992, Özyürek 2000). That speakers do this extra semiotic work to direct their representational bodily actions to recipients suggests that they assume their addressees' ability to interpret and make use of those particular actions for the at-hand task of interpretation. Gestures do more than simply make information available to onlookers, since speakers are visibly **working to make it available** by their efforts to orient to addressees (Goodwin 2000a). Speakers in the Lao kinship examples orient their bodies and their gestures to their interlocutors. They consistently place gestures in the ideal location for joint attention. Their deployment of gesture is closely coordinated temporally with associated expressions in the accompanying speech. They make extensive use of eye gaze both for directing attention to the representational structures being established and for checking receipt of the information so far supplied. There is evidence that speakers' gaze fixations on their own gestures play an important role in recipient uptake of gesturally expressed information: when a speaker looks at his or her own gestures, recipients are more likely to show evidence of having taken on board the information expressed in those gestures (Gullberg 2003).

Recipients can accordingly be observed to orient physically to the gestural contributions of speakers. An illustration of the way in which multiple individuals can physically orient to a single gestural representation comes from an interview in which two speakers are describing different kinds of fish-trapping gear used in rural Laos (see Chapter 5, above). The topic in the section of interest is the *lòòp4*, a horizontal cylinder trap (Claridge *et al.* 1997: 27). One way of using this trap is to set up a means of channelling flowing water into the trap's opening so that fish will go down into it and get stuck. The younger man on the right of the image in Figure 6.27 is describing the placement of the channelling walls which serve this purpose.

- (17) *pên3 cang1 sii4 dêj2 faa3-maaj3*
 COP like so FAC.NEWS channelling-wall
 '(They're) like this, you see, the channelling walls_{Fig. 6.27}.'

The speaker has his arms and hands placed on the table in front of him as an explicit means of illustrating the concept he is trying to convey. His arms depict the channelling walls. His splayed hands represent the channel opening up again on the downstream side. As he talks, his gaze is fixed on the space between his arms, which corresponds to the spot where the *loop4* trap would be placed. He is making an indexical gesture with his gaze in the absence of a third hand.⁹ While this 'channelling walls' representation differs from the

⁹ Here the speaker's most versatile representational resources, his two arms and hands, are occupied. Gaze becomes especially useful here (cf. Chapter 3). See Chapter 7 for a description of further techniques for dealing with representational limitations of this kind.



Figure 6.27 ‘(They’re) like this, you see, the channelling walls.’

above kinship diagrams in directly depicting the shape and spatial layout of concrete objects rather than abstract relations, the two types of representation are the same in one important respect. Their interpretation involves sustained conceptual projection of structure onto an unstructured spatial medium. In Figure 6.27, the speaker’s arms are doing the representational work. The table provides a backdrop of undifferentiated space, but like the pieces of air in which the kin diagrams are inscribed, it inherits representational structure from the speech and gesture to which it is compositionally bound.

It is in this space between the speaker’s arms that the *lòðp4* fish trap is placed. This remains implicit at the segment of discourse depicted in Figure 6.27, since the current speaker has not yet explicitly mentioned the trap. The second speaker now chimes in and makes this point explicit, providing the third arm that the first speaker is lacking. The second speaker’s left hand comes forward and taps the space on the table between the first speaker’s arms. The first speaker has assigned referential correspondence between this space and the space in the water where the trap should go (Figure 6.28).

- (18) *qaw3 lòðp4 saj1 nii4*
 take lòðp4 put HERE
 ‘The *lòðp4* is put here_{Fig. 6.28}.’

This is a typical but nevertheless extraordinary case of distributed representation and, indeed, of distributed cognition. The two men collaborate in physically and cognitively orienting to as well as constituting the diagram as a unitary composite representation. The normal situation is for representation to



Figure 6.28 ‘The *lòòp4* is put here.’

be distributed across sub-systems within a single speaker’s suite of available representational resources, including verbal and constructional linguistic categories, prosody, timing, and the myriad of visible signals including bodily orientation, gaze, and gestural tools for representing information. Distribution of information across these modes of representation lowers the overall load on any one of them (Goldin-Meadow 2003a). This becomes especially apparent when constraints are imposed on one or other mode of representation. When others step in to compensate for those constraints, as shown in Figure 6.28, both cognition and representation are distributed **across** individuals.

The phenomenon of distributed cognition is best known from cases in which the relevant representational constraints arise from the presence of a technology which operates on a larger scale than a single human being can handle (Hutchins 1995, Goodwin 1996). Such constraints do, however, also routinely occur at the smaller scale of face-to-face interaction. And when an individual’s resources for representation are heavily constrained, the problem is naturally dealt with by distributing the component representational tasks across interlocutors. This is dramatically visible in Goodwin’s research on day-to-day family interaction with Chil, a man with severe nonfluent aphasia (Goodwin 2000b, 2003, 2006, Goodwin *et al.* 2002). Goodwin writes of Chil, ‘His gestures do not stand alone, but instead count as meaningful actions by functioning as components of a distributed process in which he creatively makes use of the language of others’ (2000b: 84). This is possible because of the public orientation to communicative bodily movements and their

meanings. The example illustrated in Figures 6.27 and 6.28 shows clearly the sense in which these diagrams are both made public and publicly made for the cognitive task of articulating meaning in communicative moves. This makes them cognitive artefacts to the extent that they provide the mediating artefactual structure that enables distributed cognition (Hutchins and Hazlehurst 1995).

A second main source of evidence for regarding the communicative body as a cognitive artefact comes from cases in which speakers oblige addressees to rely on the virtual diagram for successful interpretation of the accompanying spoken utterance. This is most notable when the linguistic code features symbolic indexical expressions, which depend on something external to the expression itself for their interpretation (see Part I, this volume). The simplest type of example is the use of the body as a measure. Figures 6.29a–b each show a speaker drawing attention to a part of his body (with gaze fixed on the relevant part) to exemplify some feature of what he is saying.

In these examples, the speaker is treating a part of his body like an object linked to speech by indexical signals. But in the kinship-gesture diagrams, by contrast, the body is the pointer, not the thing pointed at. The structure of the target ‘object’ itself is manifest in empty space. There are multiple cases in the Lao kinship examples, above, in which speakers use indexical elements to refer to points in space whose reference had previously been established but which have been temporarily abandoned and unattended (e.g. Examples 3 in section 6.5.3, above, and 8 in section 6.5.8, above). In these cases, there is no perceptual evidence available to participants of any link between signifier and signified. The only way to recover reference is to have maintained a



Figure 6.29a Speaker (centre) showing the size of a species of fish with indexical reference to his little finger.



Figure 6.29b Speaker (left) showing the length of a species of fish, with reference to a point along his left arm (as indicated by the right hand gesture).

conceptually accessible representation of the set of relations established in the (invisible!) diagram. That this is actually the case is supported by participants' evident **assumption that** it is the case in certain sequences in which speakers oblige their recipients to consult the virtual diagrams for interpretation of their utterances. None of these speakers act as if they had any trouble or expected their addressees to have any trouble in keeping track of reference. This referential dependence on gesture diagrams reveals their role as cognitive artefacts – externally anchored representations which figure critically in the cognitive maintenance of, and operation upon, information.

A third basis for maintaining that the body is a cognitive artefact is an outcome of the fact that communicative bodily movements occur as elements of composite utterances. Cognitive artefacts are devices which semiotically facilitate tasks including the solving of mathematical problems, the derivation of logical inferences, the opening of doors, the operation of video cassette recorders, and so on (Norman 1988). If, in the course of solving a problem, a person carries out some operations on perceptible structures and these operations are unequivocally part of the attempt to solve that problem, then those perceptible structures are cognitive artefacts – in Norman's terms, devices 'designed to maintain, display, or operate upon information in order to serve a representational function'. Speech is a cognitive artefact in this sense. Now, if speech occurs within composite utterances in which it supports gesture and is supported by it (Hutchins and Palen 1993: 38), then gestures are cognitive artefacts for the same reason that speech is. This does not mean

that just any visually accessible bodily movement made during speech is a cognitive artefact. What counts is determined by people's differential attention to bodily movements with respect to their interpretation of the tasks a speaker is engaged in (Kendon 1978, Goodwin 1986).

People are seldom in doubt as to whether a particular hand movement is intended to be part of an accompanying utterance (e.g. pointing or making an iconic gesture) or has nothing to do with it (e.g. scratching one's nose). I investigated this in follow-up interviews conducted in Laos in which I showed the kinship interviews to a number of Lao speakers and discussed with them the gesture-diagram sequences. Without exception, viewers identified the diagramming hand movements as being part of what the speaker was saying (Kendon 2004). A typical description of the gestures was 'informing signals' (*sañña2 bong1 bòk5*). One consultant explained that the speakers in these examples were 'not just speaking with the mouth' (*bòl1 phiang2 tèè1 vaw4 kap2 paak5*). It is clearly recognizable that speakers are mobilizing everything at their disposal in order to solve a cognitive and representational problem. They are working to make sure that their performance is attended to and understood by their addressees. And their bodily movements, specifically their hand movements, are recognized by viewers as part of what they are saying. Saying things is not only, or even primarily, a linguistic activity, but a process of constructing moves out of composite resources.

Both speech and hand movements in these sequences are uncontroversially taken by observers to be part of the speaker's efforts to solve the problem of articulation for comprehension. This is the third sense in which hand gestures and other communicative bodily movements are cognitive artefacts: they function recognizably as tools for expression of what a person is saying and are thereby, again, serving as externally anchored representations which figure critically in the cognitive maintenance of and operation upon information.

6.8 Representational determinism

We currently lack much sense of the extent of cultural variation in spatial representation of abstract relations in kinship. Nor do we know much about culture-specific conventions of the use of hand gestures and other bodily movements in the spatial representation of abstract ideas (but cf. Efron 1972 [1941], Calbris 1990, Wilkins 1997, Kita *et al.* 2001, *inter alia*). There is much at stake if patterns turn out to be significantly different across ethnographic settings. Zhang (1997: 213) proposes a representational determinism based on the observation that 'different representations of a common abstract structure can cause dramatically different cognitive behaviors' (cf. Norman 1991: 34, Wilkins 2002). Recent neo-Whorfian approaches to the language-cognition

interface (Lucy 1992a, b, Gumperz and Levinson 1996, Lee 1996, Bowerman and Levinson 2001, Levinson 2003, Gentner and Goldin-Meadow 2003) suggest ways in which this broader notion of representational determinism (or representational relativity principle) could be rigorously investigated. Effects may be expected not only in the cognitive behaviour of the person producing the representation but also in that of the person perceiving and interpreting it. In all but one of the different spatial mappings of sibling relations in the examples summarized in Figure 6.26, for example, siblings are separated laterally. The exception is Figure 6.26g (Example 8, section 6.5.8, above), where the issue at hand is not collateral lineage or equivalent status of siblings but the relative age of the two siblings concerned (the speaker's wife and her sister). When relative age is the only thing that matters in a relationship being represented, it becomes possible to use height alone as a basis for differentiation in spatial mapping. This difference in representation arises in the performance of a single speaker. What if distinctions like this were to define the habitual behaviour of a whole cultural group? Kita *et al.* (2001) document a striking difference in the use of gesture space in a comparison of hand gesture in narrative by individuals from two Mayan cultures, Mopan in Belize and Yucatec in Mexico. They write: 'For Yucatec Mayans, but not for Mopan Mayans, conceptually distinct entities can be located at different points along the projected lateral axis. Consequently, the "shape" of abstract thought is different in the two cultures: time flows and a plot develops along different axes, and contrasted entities are localized differently' (Kita *et al.* 2001: 137; cf. also Núñez and Sweetser 2006). Such studies suggest directions for the empirical ethnography of gesture and its cognitive implications.

6.9 Concluding remark

Bouquet (1996: 62) asks about 'the fate of the genealogical diagram' in kinship research. I have tried to sketch some issues that I think deserve to be on the agenda. The modern anthropological study of kinship is a pursuit of the domesticated mind (Goody 1977). A graphic representation like a scientific kinship diagram is a tool for thinking, 'a facilitating device' (e.g. Goody 1977: 109; cf. Norman 1991: 17, Hutchins 1995, Clark 2002). So, too, are the wealth of other types of perceptually accessible representation, not all of which endure the way print does. This chapter has investigated one genre of genealogical diagram which has no intrinsic or necessary relation to conventions of literate culture. It comes in the form of fleeting, evanescent sketches which speakers produce online using co-speech hand gestures. How much do we know about the visual representation of abstract kinship

relationships in societies whose members are not (diagram-)literate? What is the cognitive status of the ethnogenealogical diagram?

For daily puzzles of expressing meaning, the body is an abacus, a sextant, a pencil and paper. Hand gestures, and the human body more generally, afford the special cognitive advantages claimed to inhere in the enduring visuospatial modalities of iconography, semasiography, and print. As Hutchins and Hazlehurst (1995: 64) remind us, 'no individual can influence the internal processing of another except by putting mediating artifactual structure in the environment of the other'. The **requirement** to externalize ideas may be what is behind our propensity to use space in thinking wherever we can (witness the ubiquity of spatial metaphor in language). Even without visual diagrams, we are still using 'internal imaging processes' in order to access the inferential affordances of visuospatial representation (Larkin and Simon 1987: 92). Larkin and Simon (1987: 97) speculate that 'mental images play a role in problem solving quite analogous to the role played by external diagrams (and that this role is also played in the two memories, internal and external, in concert)'. Is the task of verbalization a matter of constructing a linear version of our two- and three-dimensional thoughts? Or, in producing gestures, are we transforming our one-dimensional thoughts by giving them visuospatial properties? Larkin and Simon (1987: 72) describe what happens when they give their experimental subjects a complex physics problem: 'Everyone we've observed reaches for pencil and paper, and draws a sketch of the situation.' Ordinary people trying to communicate their thoughts do the same, not reaching for pencil and paper but sketching with bare hands in mid-air. As Efron wrote in his classic work *Gesture, race, and culture*, gestural behaviour during speech is 'an intrinsic part of the thinking process' (1972/1941: 105). Gestures provide a primordial sketchpad for organizing thoughts and displaying them to others, a technology of both the intellect and the body, that supreme tool kit for overcoming our lack of telepathy.

Being context-bound and evanescent, sequences of gestures do not allow the 'reflection' that Goody (1977: 109) has identified as such a special affordance of written language and its relatives, but the case for a **unique** influence of representations in the printed modality may have been overstated. Oestermeier and Hesse show that with diagrammatic or graphical externalization, we can 'transform abstract relationships into visible spatial ones, and thereby inspect and control argumentative and causal relationships' (Oestermeier and Hesse 2000: 81). But they claim in addition that we can 'thereby inspect and control argumentative and causal relationships **in a way completely unknown to illiterate societies**' (ibid., emphasis added). The Lao data discussed here show that hand gestures can and do transform abstract relationships into visible spatial ones and, indeed, allow us to thereby inspect and control argumentative and causal relationships. Members of all

societies, literate or not, make hand gestures while they speak. Do they all produce the kinds of diagrams these Lao speakers produce when talking about kinship? The kinship domain is just one among many which will yield fertile data in exploration of this little-charted territory: the body as cognitive artefact.

A wave is not an individual except for the man who regards it and sees it advancing toward him.

Maurice Merleau-Ponty, 1942

This chapter zeroes in further on the elaborate enchronic structure of composite utterances in sequence, tracking the consequences of their uncanny persistence and coherence. We here highlight their extraordinary robustness and flexibility. We shall consider two examples in detail, from informal interviews about kinship and kin terminology, discussed already in Chapter 6. The diagrams not only display spatial integrity and temporal persistence but also may be adjusted, revised, or edited online, as the need arises. There are at least three reasons why speakers who create gesture diagrams or models may find themselves in need of making some revision. First, they may make errors in representation (e.g. I may have forgotten some component or I may have the proportions wrong and find myself now unable to fit some part of the structure in). Second, they may reach the physical limitations of gesture space, placing some part of their diagram at the extreme of arm's reach, only to find that the next part of their developing diagram needs to be placed even farther out (or, indeed, finding that some part of the diagram needs to be placed closer to their body than they have allowed space for). Third, the semiotic motivation for the structural configuration of the diagram may change as the diagram develops. In a purely iconic diagram, this may be due to a change in perspective (e.g. from diagrammatic to viewer perspective; Emmorey 2001), while in a diagram that uses space metaphorically in depicting some non-spatial structure, the nature of the metaphorical mapping (e.g. whether only relative height carries a meaning or whether laterality becomes significant) may change as the interaction unfolds. Each of these motivations for adjustment or revision of a gesture diagram is exemplified in the data discussed in this chapter.

7.1 Example 1 – marriage between cousins

As discussed in Chapter 6, Lao speakers report that marriage between second cousins is regarded as generally permissible, whereas marriage between first

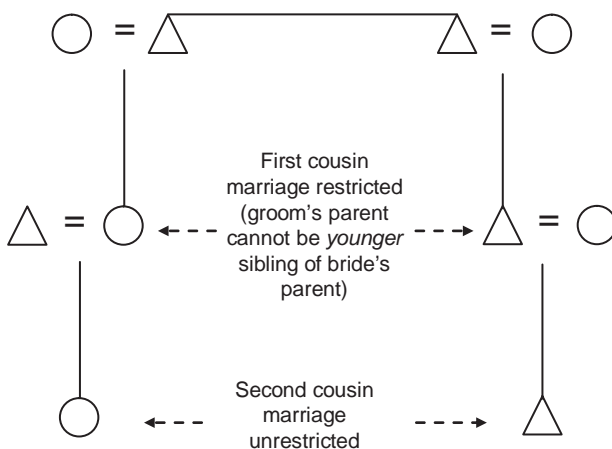


Figure 7.1 Lao speakers' reported attitudes to marriage among cousins.

cousins is regarded as permissible only under certain conditions (specifically, when the parent of the groom is the older sibling – not the younger sibling – of the parent of the bride) (Figure 7.1).

The speaker in this example (Mr Phouthong) wants to explain that it is generally permissible for second cousins to marry, but not for first cousins (as illustrated in Figure 7.1). He constructs a hypothetical scenario in which he and the man seated to his right (Mr Naak) are brothers, and they both have children (who are first cousins to each other), and these children have children in turn (who are second cousins to each other). Mr Phouthong is now concerned with depicting relations between himself and his hypothetical brother, their hypothetical children, and their hypothetical children's hypothetical children. Having invited his addressees to suppose that he is the younger brother of Mr Naak, he begins as in (1).

- (1) [khòðj5] qaw3 [mia2] mii2 [luuk4] qòðk5 maa2
 1SG.P take wife there.is child exit come
 '[I]_{Fig. 7.2a} take a [wife]_{Fig. 7.2b}, and we have a [child]_{Fig. 7.2c}.'

With his dominant (right) hand, the speaker makes three deictic gesture strokes, each coinciding with the verbal articulation of its intended referent ('I', 'wife', and 'child', respectively). These three gestures are made with relaxed, slightly curled B-hand (a hand shape in which the palm is open and flat, with fingers not spaced apart; McNeill 1992: 87–88), each gesture slightly contralateral to the speaker's front left, the first slightly to the side with bent arm (as shown in Figure 7.2a), the second straightened to be more



Figure 7.2 (a) ‘I . . .’; (b) ‘. . . take a wife . . .’; (c) ‘. . . and we have a child’.

central but tilted slightly downward (Figure 7.2b), and the third raised and fully extended to the front (Figure 7.2c).

With this series of abstract pointing gestures, the speaker has drawn a line of descent, emanating straight out from his own body, pointing forward and slightly to his left.

Next, he turns his attention and gaze to Mr Naak and almost exactly repeats the previous move, but now with reference to Mr Naak, his older brother.

- (2) *laø qajø-[naak4] qaw3 [mia2] mii2 [luuk4] qòðk5 maa2*
 PRF EB-N. take wife there.is child exit come
 ‘[Naak]_{Fig. 7.3a} takes a [wife]_{Fig. 7.3b}, and they have a [child]_{Fig. 7.3c}.’

The three gesture strokes marked by square brackets in Example (2) are illustrated in Figures 7.3a, 7.3b, and 7.3c, respectively.

Two things can be noted. First, the speaker’s hand shape in Figures 7.2a–7.3c remains more or less the same – a relaxed, slightly curled B-hand. Second, the speaker accompanies his abstract pointing gestures with eye gaze, reinforcing the deictic function of the gestures. Gazing at a manually indicated chunk of space puts extra attentional focus on it as a conversational referent, doubling the baptismal indexicality (Haviland 2000: 20). Here the speaker is not gazing at his gesture, but farther away, so that the relevant chunk of space is located at the point of intersection of the vectors of his gaze and pointing gesture.¹

¹ This is how reference-activated chunks of space can be set at distances away from the body, which would be impossible with only a single vector. It also shows how a token space can be set up outside the proximal gesture space (*pace* Liddell 1995: 33).



Figure 7.3 (a) '[Naak] ...'; (b) '... takes a [wife] ...'; (c) '... and they have a [child]'.

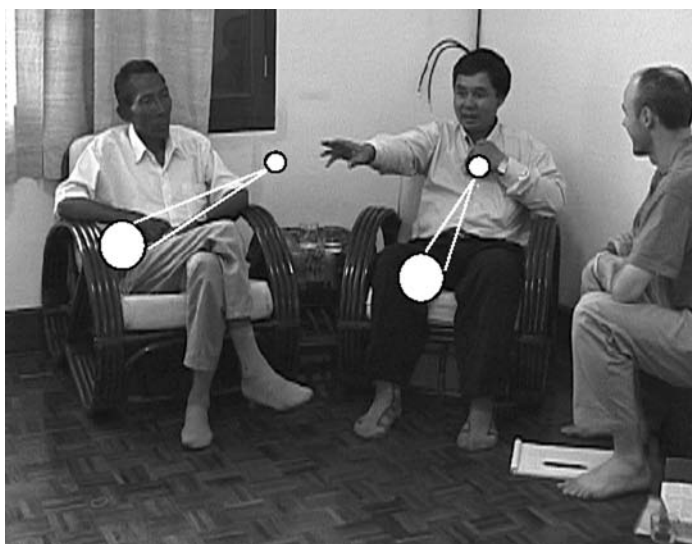


Figure 7.4 Diagram of two brothers and their respective children.

The speaker has now set up two parallel lines of descent, emanating forward and away from himself and his brother, respectively, on a more or less flat plane. The important nodes on the lines of descent at this moment are the two brothers and their respective children. The children are denoted by the gestures shown in Figures 7.2c and 7.3c, both of which involve the right arm fully outstretched. The current state of Mr Phouthong's diagram can be depicted as in Figure 7.4 (in which the two smaller white circles represent the speaker and his hypothetical brother, the two larger circles depict their respective children, and the connecting white lines represent the father–son relationship).



Figure 7.5 (a) ‘If [Naak’s child] is a male . . .’; (b) ‘. . . and if [my child] is a female, then they can marry’.

The speaker now explains that the two children can marry only if his own child (i.e. the child of the younger of the two siblings) is the bride, not the groom. As he mentions the respective children, he points (with index finger and gaze) to the chunks of space that have just been set up to refer to these referents (i.e. the larger white circles in Figure 7.4). These two finger-and-gaze pointing gestures are illustrated in Figures 7.5a and 7.5b.

- (3) *khan2 [luuk4 qajø-naak4] hanø pên3 phuø-saaj2*
 if child eB-N. TPC.DIST be CT.HUM-male
khan2 [luuk4 khòoj5] pên3 phuø-ñing2 niø mèn1 qaw3
 if child 1SG.P be CT.HUM-female TPC be marry
daj4
 can

‘If [Naak’s child]_{Fig. 7.5a} is a male and if [my child]_{Fig. 7.5b} is a female, then they can marry.’

The hand shape is now an index-finger pointing gesture (relaxed G-hand, a hand shape in which the index finger is sticking out straight and all other fingers are curled over; McNeill 1992: 87–88), and, once again, gaze is aligned with these pointing gestures, such that looking and pointing are directed to the same space.



Figure 7.6 (a) Hand relaxed in lap, prior to speaking; (b) ‘So, now to [calculate] the further connections.’

So far this is a typical use of abstract pointing to activate chunks of space with referential association to specific referents. Here, the speaker has set up no less than four such chunks (see Figure 7.4), and his deictic gestures in Figures 7.5a and 7.5b (accompanied by the speech in (3)) demonstrate that the overall structure has cohesion and temporal persistence. That is, having set up a node referring to his own child in Figure 7.2c, and having subsequently made gestures to more distant locations, he is able to point back to the same chunk of space in Figure 7.5b (more than seven seconds later).

Now, in order to continue his hypothetical outline, the speaker wants to bring the next generation of the family tree into the discussion, namely the children of the two cousins he has just established. He begins with hands relaxed in his lap, as illustrated in Figure 7.6a.

He begins the next phase of his explanation with a metanarrative comment, explicitly announcing that he will now move on with his calculations.

- (4) *lèèw4 laø [nap1-thùù3] tòø-tòø1 paj3 baat5-niø*
 finish PRF calculate RDP.A-connect go THZR
 ‘So, now to [calculate]_{Fig. 7.6b} the further connections.’

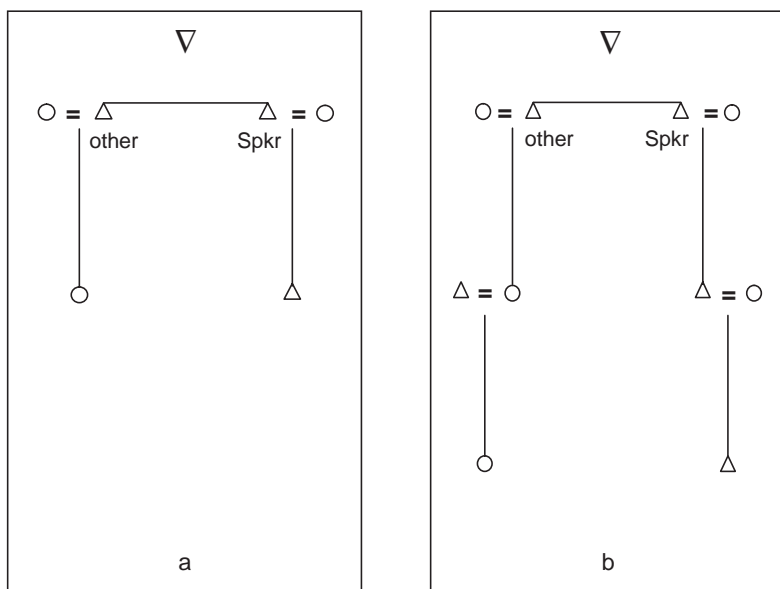


Figure 7.7 (a) First generation of descent from speaker and speaker's brother; (b) First and second generations of descent from speaker and speaker's brother.

The gesture accompanying this comment is a simultaneous raising of the two hands followed by straightening of the elbows until the arms are forward and outstretched (Figure 7.6b). He is representing the two children of the first generation (first introduced in Figures 7.2c and 7.3c and marked as large white circles in Figure 7.4), one on either side.

Here the speaker encounters a representational problem. He has so far created a diagram depicting the first generation below himself and his brother. This is shown in Figure 7.4 and more abstractly in Figure 7.7a. His problem now is to add to this diagram the next generation – that is, the children of his own children as well as the children of his brother's children. He needs to extend the present diagram so as to produce Figure 7.7b.

Given the speaker's current physical orientation and the logic of his developing diagram, he is unable to proceed without making adjustments. His fully outstretched arms (shown in Figure 7.6b) indicate the first set of children. The diagram is motivated by a metaphor that maps generations of descent onto steps forward and away, and since his arms are already fully outstretched, he is lacking the necessary forward reach to represent the next

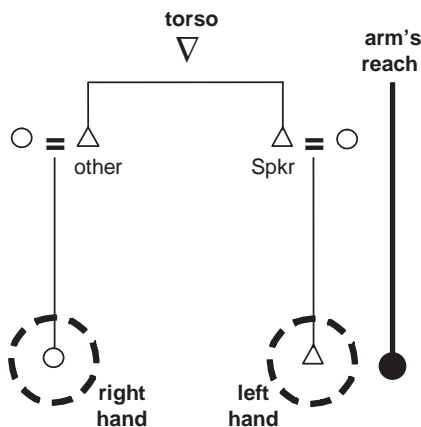


Figure 7.8 Mapping of speaker's gesture (Figure 7.6) onto diagram of first generation of descent (Figure 7.7a).

generation. This situation is depicted in Figure 7.8, a schematic representation of Mr Phouthong's physical relationship to the diagram he is working on, as it is at the stage depicted in Figure 7.6b and Figure 7.7a.

The speaker solves this problem with a combination of two devices. Beginning from the position shown in Figure 7.8, he experiences disfluency, as he slightly bends his elbows and brings his dominant (right) hand forward and across with his left hand held in space, relaxing only slightly.

- (5) [lang3-caak5] ...
 after
 '[After]_{Fig. 7.9a} ... (1.8s)'

Next, Mr Phouthong performs a hold-and-drag move with his left hand, bringing the first-generation chunk of space, anchored to his left hand, closer in towards his own body. This action has already begun in Figure 7.9a and is finished by the time the speaker mentions the third generation. By Figure 7.9c, the speaker's left hand is now resting on his knee. At the same time, he moves his torso forward (compare Figures 7.9a, 7.9b, and 7.9c), increasing his outward reach into space. When he moves forward, the diagram does not automatically move with him, despite the fact that it was first established in relation to his torso position. This is visible in that the speaker appears to be physically leaning forward and over the diagram. Once established, the diagram maintains more or less stable coordinates in absolute space. By the time Mr Phouthong



Figure 7.9 (a) '[After] . . . (1.8s)'; (b) 'I take a wife and a child. [My child] takes a wife . . .'; (c) '. . . and has [a child] too'.

makes explicit reference to his grandchild in the diagram, his body is far forward and his right arm is once again fully outstretched (Figure 7.9c).

- (6) *khòdɔ̌j5 qaw3 mia2 mi2 luuk4 – [luuk4 khòdɔ̌j5] kaø-phaɔ*
 1SG.P take wife have child child 1SG.P T.LNK-CONTR
paɔj qaw3 mia2 qiik5 phaɔ mi2 [luuk4] qiik5 nòq1
 go take wife more CONTR have child more QPLR.AGREE
 'I take a wife and have a child. [My child]_{Fig. 7.9b} takes a wife and has [a child]_{Fig. 7.9c} too.'

This second move can be illustrated more schematically as in Figure 7.10 (as a second step from Figure 7.8, showing the transition from Figure 7.7a to Figure 7.7b).

The adjustments that the speaker has made to his own side of the diagram have resulted in there being now not one but two chunks of space positioned in a line on the sagittal axis, standing for the two descending generations, respectively. He now explicitly creates the same adjustment on Mr Naak's side of the diagram, but assuming a symmetry to the diagram, he no longer needs to hold the first generation chunk of space and drag it closer in to Mr Naak's body. He merely uses abstract pointing gestures to create the relevant two distinct chunks of space referring to the further generations. Mr Phouthong almost exactly repeats the structure of what he said in Example (6), only now referring to Mr Naak.

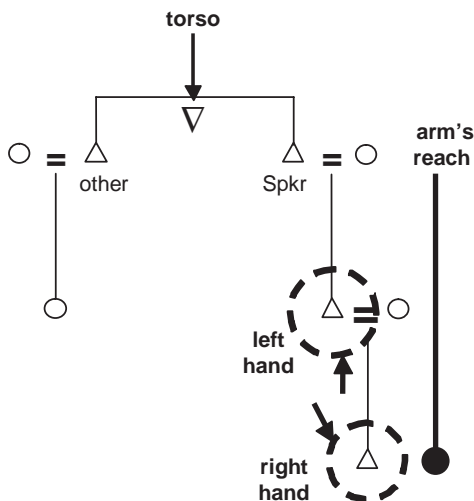


Figure 7.10 Mapping of speaker's sequence of gestures (Figures 7.9a and 7.9b) onto partial diagram of first and second generations of descent (cf. Figure 7.7b).

- (7) *qajø-naak4 qaw3 mia2 mii2 [luuk4] – luuk4 qajø-naak4*
 eB-N. take wife have child child eB-N.
hanø phatø qaw3 mia2 phatø mii2 [luuk4] qiik5
 TPC.DIST CONTR take wife CONTR have child more
 'Naak takes a wife and has [a child]_{Fig. 7.11a}. Naak's child takes a wife
 and has [a child]_{Fig. 7.11b} too.'

The speaker is now leaning visibly farther forward than he was in the earlier sequence shown in Figures 7.3a, 7.3b, and 7.3c, increasing his reach into gesture space. Although these chunks of space are abstract referents, or in Liddell's terminology token spaces, they are located not in the immediate gesture space, but beyond it (*pace* Liddell 1995: 33). Mr Phouthong has now created a virtual diagram of the form shown in Figure 7.12. (In this diagram, (N↓) and (P↓) represent the viewpoints of Mr Naak and Mr Phouthong from where they are sitting, boldface letters refer to nodes on the diagram (N = Mr Naak, P = Mr Phouthong), and subscript numbers refer to generations of the men's children.)

The gesture diagram itself is different from Figure 7.12 in one important way, namely that whereas the artefactual signifying material in Figure 7.12 is



Figure 7.11 (a) ‘Naak takes a wife and has [a child]. Naak’s child takes a wife . . .’; (b) ‘. . . and has [a child] too’.

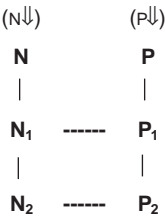


Figure 7.12 Graphic depiction of a gesture diagram representing first and second generations of descent from speaker and speaker’s brother (cf. [Figure 7.7b](#), [Figure 7.13](#)).

ink on the page, there is no artefactual signifying material corresponding to the gesture diagram illustrated in [Figures 7.2a–7.11b](#). The hand is a visible artefactual signifier, but it is not the diagram. The hand is not analogous in this case to ink on the page, but to the pen that inscribes it. Further, when the hand does signify in a manner analogous to ink on paper (e.g. when serving a direct modelling function; cf. [Chapter 5](#), above), it does so only evanescently. Unlike ink on paper, it does not endure, and the diagram thus has an inescapable temporality. Although the diagram is wholly in the mind of the interlocutors involved, as a structured set of semiotic associations between chunks of space and conceived referents, it nevertheless has structural

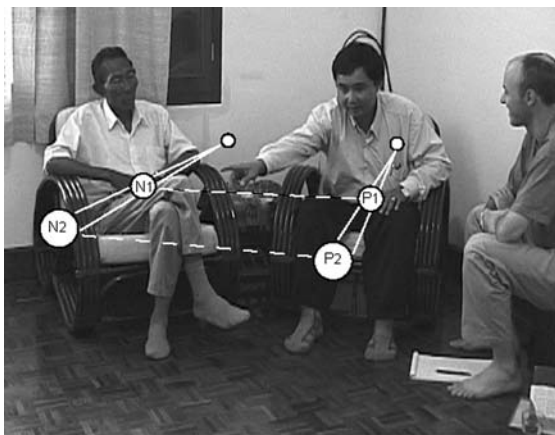


Figure 7.13 Diagram of two brothers and their respective children and grandchildren, focusing on relations between cousins, and between second cousins.

integrity and temporal persistence. Figure 7.13 offers an illustration of the diagram as it would be projected by the interlocutors in the actual space.

Now that Mr Phouthong has established this structure, he can continue his exposition while free to indicate the various nodes with spatially oriented deictic gestures as he makes reference to different individuals concerned.

He now states that marriage between second cousins (i.e. between N2 and P2 in Figure 7.12) is permissible. Throughout the utterance his arms are outstretched, index fingers deictically referring to the two grandchildren, while his hands are panning repeatedly, outward to the sides (Figure 7.14a) and inward to the centre (Figure 7.14b) in a representation of their coming together in marriage. The hands move as if tracing back and forth along the dotted line running between the two largest white dots in Figure 7.13. The gesture overlaps fully with the utterance:

- (8) *qaw3 kan3 baat5-niø phuak4 nan4 naø –*
 take COLL THZR group DEM.NONPROX TPC.PERIPH
juu1 phun4 khaw3 cang1 qaw3 kan3 daj4 bèèp5 man2
 be.at DEM.FAR 3PL.B thus take COLL can like 3.B
kaj3 dêê4
 far FAC.FILLIN
 ‘(They) get married, those ones – being yonder, they can marry, since it’s far.’

It is probably impossible to say whether this gesture employs tracing (in which a line between N2 and P2 is drawn – cf. Figure 7.13 – representing the



Figure 7.14 (a) ‘(They) get married, those ones . . .’; (b) ‘. . . being yonder, they can marry, since it’s far’.

link between them) or an analogic enactment (of the event of their coming together). In either case, however, both the tracing and the enactment are played out at a distance, beyond the gesture space, projected by the pointing index fingers. This can be seen by the fact that the speaker’s eye gaze appears not to be on his fingertips but at points in the distance.

Next, Mr Phouthong remarks that the marriage of first cousins is not preferred. In doing so, he draws his hands in closer to his body, bending his elbows and angling the hands to point down (onto the flat plane of the gesture diagram), deictically referring now to N1 and P1 in Figure 7.13. The greater proximity of these two nodes is reinforced by his gaze behaviour. Rather than gazing out beyond where his hands are (as he does in Figures 7.14a and 7.14b), he now looks down onto his hands, first at N1 (Mr Naak’s child), then at P1 (his own child), as shown in Figures 7.15a and 7.15b, respectively.

- (9) *[khan2] vaa1 juu1 [thitø-thit2] kan3 niø phatø*
 if COMP be.at RDP.A-close COLL TPC CONTR –
khaw3 phatø bõø haj5 qaw3
 3.B CONTR NEG allow marry
 ‘[If]_{Fig. 7.15a} they’re close [together]_{Fig. 7.15b}, like so, they don’t allow them to marry.’



Figure 7.15 (a) ‘[If] they’re close [together] . . .’; (b) ‘. . . like so, they don’t allow them to marry’.

In both Examples (8) and (9) the speaker uses a spatial metaphor in speech as well as gesture, using terms for spatial distance (i.e. ‘far’ and ‘close’) to express kinship relations.

Next, Mr Phouthong is able to exploit the presence of the diagram for further comment on the use of the kinship term *laan3* ‘grandchild, nephew, niece, cousin via parent’s younger sibling’ (see Chapter 6). He first refers back to the two grandchildren.

- (10) *khan2 vaa1 – sommut1 vaa1 [pên3 laan3] niø . . .*
 if COMP suppose COMP be *laan3* TPC
 ‘If – suppose they [are *laan3*]_{Fig. 7.16a} . . .’

Earlier in the conversation, the term *laan3* had already been defined as ‘the child of one’s own child’, and in Figure 7.16a Mr Phouthong is referring by his gesture to the two grandchildren who would be referred to as *laan3* in this sense by both men (i.e. N2 and P2 in Figure 7.13). His hands are fully outstretched, pointing well forward. Mr Phouthong now goes on to add that the two men would also each use the term *laan3* to refer to the child of the other (i.e. to the child of one’s own brother).

- (11) [*khòøj5*] hòòng4 [luuk4] qajø-naak4
 1SG.P call child eB-N.
kaø ñang2 pên3 laan3 juu2
 T.LNK still be *laan3* FAC.WEAK
 ‘[I]_{Fig. 7.16b} still call Naak’s [children]_{Fig. 7.16c} *laan3*, nevertheless’.



Figure 7.16 (a) ‘If – suppose the [are *laan3*] . . .’; (b) ‘[I] still call . . .’; (c) ‘. . . Naak’s [children] *laan3*, nevertheless’.



Figure 7.17 (a) ‘(and) [Naak] calls . . .’; (b) ‘. . . [my children] *laan3*, too’.

In making this utterance, he first index-finger-points to his own chest (referring to ‘I’, Figure 7.16b), and second to Naak’s child, with index-finger-point and gaze directed at the node N1 of Figure 7.13 (as illustrated in Figure 7.16c).

Following this, he states the converse, pointing first to N (Figure 7.17a) and second to P1 (Figure 7.17b) on his diagram (cf. Figure 7.13).

- (12) [qajɔ-naak4] hòòng4 [luuk4 khòòj5] kaø pên3 laan3 khùù2
 eB-N. call child 1SG.P T.LNK be laan3 like
 kan3
 COLL

‘(and) [Naak]_{Fig. 7.17a} calls [my children]_{Fig. 7.17b} *laan3*, too.’

The pointing gesture shown in Figure 7.17b indicates the node P1. Greater laterality and downwardness in angle of the outstretched-arm index-finger pointing gesture indicates that the node is closer to the speaker’s body. In addition, the speaker’s fixation of gaze on the point, looking at the relevant chunk of space as if he were looking at a thing, puts greater focus on it as an indexically created referent.

7.2 Discussion of example 1

In this example, the speaker utilizes not only his own gesture space, but also the gesture space of his interlocutor, borrowing it for the purpose of his exposition. Unlike the illustrations in Figures 7.1 and 7.12, Mr Phouthong’s diagram relies on the mental projection of his interlocutor for it to work, indeed for it to exist. Ink-on-paper illustrations also require a cognitive contribution in order for them to signify, but in the case of the gesture diagram no physically discrete signifier need persist. Mr Phouthong makes heavy use of abstract pointing gestures, coordinated closely with speech and gaze, to activate distinct chunks of space, assigning each one reference to a distinct but related idea, and the virtual structure he creates remains in place for some time. The diagram has spatial cohesion and temporal persistence, especially clear in that points on it become referents for later deictic gestures. The speaker successfully spatializes and makes visually accessible an inherently non-spatial and invisible structure.

Because the production of gesture diagrams is an enchronic, multi-move process with a necessarily temporal dimension, there is always the possibility of running into problems in construction of the diagram. The above example illustrates one such case, and we have observed one speaker’s techniques for overcoming the problem of having painted himself into a corner. The moment when Mr Phouthong has difficulties in Example 1 is illustrated in Figures 7.6b–7.9c (and accompanying Examples (5) and (6)). He has run out of gesture space and has not foreseen the constraints that his diagram presents. His initial move of setting up chunk-of-space reference points for the first generation (the two first cousins) with full arm extension to the front results in a problem when it comes time to add the next generation out. He has no more forward extension physically available (cf. Figure 7.8). He needs to spatially differentiate his already established first set of children from the next generation, and this means bringing in the former, closer to himself, and

establishing the latter in a space further out (cf. Figure 7.10). He achieves both by combining the following three editing techniques:

- (i) *hold-and-drag*
hold a reference-activated chunk of space and drag it to a preferred location in the gesture space (in this case, closer to the body; Figures 7.6b–7.10);
- (ii) *hold-and-work-with-free-hand*
hold a reference-activated chunk of space while the free hand works on an extension or redesign of some part of the diagram (Figures 7.6b–7.10);
- (iii) *move-body-into-new-space*
move the body (in any direction), allowing the hands to reach into new space not hitherto accessible (compare Figures 7.9a and 7.9b); meanwhile, the diagram maintains absolute spatial coordinates and does not shift with the movement of the speaker's body.

This completes the description of Example 1. Our next example concerns kin terminology used between siblings.

7.3 Example 2 – use of kin terms between siblings

In this example, Mr Naak is discussing the meaning of the two terms *qaaj4* ‘older brother’ and *nòong4* ‘younger brother’. He will eventually depict a structure along the lines of Figure 7.18.

I will henceforth use B1, B2, B3, and B4 to refer to these individuals. The speaker is B2.

The speaker states that he himself has one older brother, for whom he uses the term *qaaj4* ‘older brother’. When asked to define the term *qaaj4*, he first states that the term refers to ‘a male who is born before (you)’. This definition is too broad (e.g. it would be true for *phòòl* ‘father’), and so he then sets out

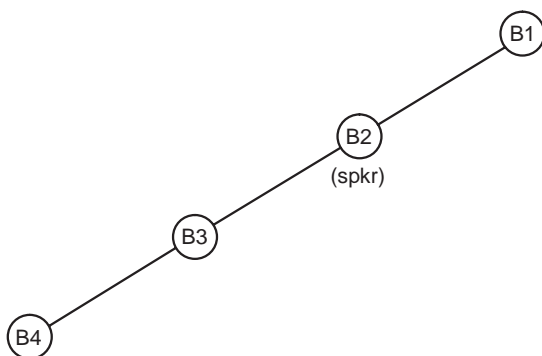


Figure 7.18 Schematic diagram of the relationship between four brothers, numbered 1 to 4 in order of birth.



Figure 7.19 (a) ‘(I mean that it’s the [father] . . .’; (b) ‘. . . of [me] . . .’; (c) ‘This one here – is the one who brings about the creation – is the one who brings about the [creation] (of) the . . .’; (d) ‘. . . this [qaaj4] (the oldest brother)’.

to differentiate his father from his older brother. He first sets up with his gesture and gaze a reference point for his father, index-finger pointing to a space in the lower portion of gesture space immediately in front (Figure 7.19a). Second, he points to himself by turning his index-finger point slightly back in toward his body (Figure 7.19b).

- (13) *maaj3-khuam2 vaa1 mèn1 [phòò1] khoòng3 [haw2] niø*
 mean COMP be father of 1.FA TPC
 ‘(I mean that it’s the [father]_{Fig7.19a} of [me]_{Fig. 7.19b} . . .’

The gesture referring to ‘father’ (Figure 7.19a) is reinforced by gaze toward the referent chunk of space. Mr Naak immediately continues, with ‘father’ as topic, introducing the first son of the father (i.e. the speaker’s older brother B1).

- (14) *phuø-nii4 niø – pèn3 phuu5 haj5 kamneet5 – pèn3 phuu5*
 MC.HUM-DEM TPC be person give be.born be person
haj5 [kamneet5] phuø – phuø [qaaj4] niø keet5 maa2
 give be.born CT.HUM CT.HUM qaaj4 TPC born come
 ‘This one here – is the one who brings about the creation – is the one who brings about the [creation]_{Fig. 7.19c} (of) the – this [qaaj4]_{Fig. 7.19d} (the oldest brother).’

The gestures of Mr Naak’s right hand are complex, and here I note just two clear points in the sequence, marked with square brackets in (14). First, a metaphoric/deictic gesture corresponds to the father’s bringing about the creation of the first son, with downward-pointing basket hand depicting creation, as Figure 7.19c. Second, a clear index-finger pointing gesture (Figure 7.19d) activates a chunk of space low in the speaker’s gesture space, directly in front, as a reference point for the first-born son (the speaker’s older brother).



Figure 7.20 Diagram of a father and son.

The move illustrated in Figures 7.19c and 7.19d metaphorically maps descent from father to son onto a vertical plane. This is in contrast to the horizontal away-going mapping selected by Mr Phouthong in Example 1, above. It can be illustrated as in Figure 7.20, with white circles representing the father (Fa) above, and the son (B1) below, with a vertical white line between them indicating their relationship.

Throughout the rest of this example, however, Mr Naak is no longer concerned with the relation between father and son and does not mention the father again. From this point, he is concerned exclusively with the relation between siblings. In particular, he wants to explain that the *qaaj4-nòng4* ‘older brother–younger sibling’ relationship is non-reciprocal. Thus, Mr Naak calls his older brother *qaaj4* ‘older brother’, whereas he is called *nòng4* ‘younger sibling’ in return. In the next part of the sequence, Mr Naak reiterates a point he has already made, namely that the first-born male is the *qaaj4* ‘older brother’ to all others, and as he finishes this statement, he uses a thumb-and-forefinger-touching gesture, pointing up (shown in Figure 7.21a) as if presenting the older brother for his addressee’s consideration.

- (15) ... [cùng1] *qeen4 vaa1 pèn3 qaaj4 daj4*
 so call COMP be qaaj4 can
 ‘... [thus]_{Fig7.21a} we can call him *qaaj4*.’

He now states his own position in the sibling ranking, which, as established earlier in the interview, is second brother.



Figure 7.21 (a) ‘... [thus] we can call him *qaaj4*’; (b) ‘[I] am the ...’; (c) ‘... [*nòòng4*], born second’.

- (16) [haw2] *laø pên3 [nòòng4] keet5 phuø thii1 sòòng3*
 1.FA PRF be younger.sibling born MC.HUM ORD two
 ‘[I]_{Fig. 7.21b} am the [*nòòng4*]_{Fig. 7.21c}, born second.’

The speaker first points to himself, corresponding to ‘I’ (Figure 7.21b) and then points to a chunk of space low and directly in front (Figure 7.21c), referring to ‘the younger brother’. Both these points refer to the speaker himself. The gesture in Figure 7.21b is oriented toward himself in the speech event, whereas the gesture in Figure 7.21c is oriented toward himself in the diagram (or the ‘narrated event’; Jakobson 1971). Now the speaker is embarking on a diagram depicting the relationships between brothers ranked by relative age (see Figure 7.18). The next thing Mr Naak does is to bring in a third brother, his own *nòòng4* ‘younger sibling’. This is achieved by contrast with the reference point he has just created in the low front gesture space (referring to himself as second brother in Figure 7.21c). Already pointing at the space he has created for reference to himself as the second brother in Figure 7.21c, he now makes a slight but very sharp shift to his right (reinforced by gaze) in establishing a fresh and distinct chunk of space to refer to the third brother.

- (17) *baat5-niø phuø-nòòng4 [phuø-nan4] laø*
 THZR MC.HUM- *nòòng4* MC.HUM-DEM.NONPROX PRF
mènè1 pên3 ... phuø-thii1 keet5 lun4 lang3 khòòng3
 be be MC.HUM-REL born series after of
khòòj5 qiik5 baat5-niø
 1SG.P more THZR
 ‘Now, [that one]_{Fig. 7.22a} who is (my) *nòòng4* is the one who ... the one who is born after me in addition.’

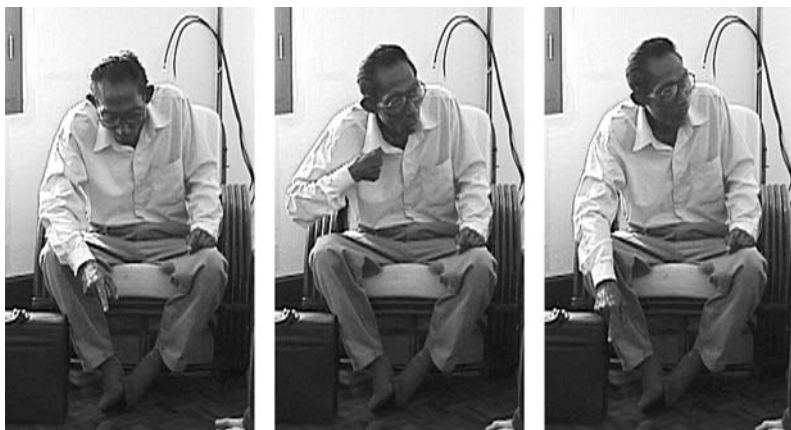


Figure 7.22 (a) ‘Now, [that one] who is (my) *nòng4* is the one who . . . the one who is born after me in addition’; (b) ‘Now, [I] was born . . .’; (c) ‘. . . before [that one] – that one has to call me *qaaj4*.’

This is followed by a comment on how he, as second brother, will address this third brother, and how the third brother will address him in return.

- (18) *baat5-niø* [*khòj5*] *keet5* *kònn1* [*phuu5-nan4*] –
 THZR 1SG.P born before MC.HUM-DEM.NONPROX
phuø-nan4 *tòng4* *qeen4* *khòj5* *pên3* *qaaj4*
 MC.HUM-DEM.NONPROX OBLIG call 1SG.P be *qaaj4*
qiik5
 more
 ‘Now, [I]_{Fig. 7.22b} was born before [that one]_{Fig. 7.22c} – that one
 has to call me *qaaj4*.’

Here, Mr Naak uses pronounced pointing gestures to single out the two referents, this time pointing to himself when referring to himself (Figure 7.22b) and pointing to the newly established chunk of space, low and slightly to the right of centre when referring to the third brother (slightly editing the diagram by pushing the chunk of space farther out in Figure 7.22c and differentiating it more clearly).

Here, as in so many examples we have seen so far, the timing of speech and gesture is closely coordinated. The speaker’s mentions of different referents occur at the same time as the stroke of deictic gestures referring to the same referents. This close link in timing is a default indexical cue for associating hand movements and speech in composite utterances.

Now Mr Naak encounters a representational problem. He wants to explain that as a middle brother, he is *nòng4* ‘younger sibling’ with respect to his older brother, yet *qaaj4* ‘older brother’ with respect to his younger brother. To make this point, he now wants to refer once again to the first brother, who

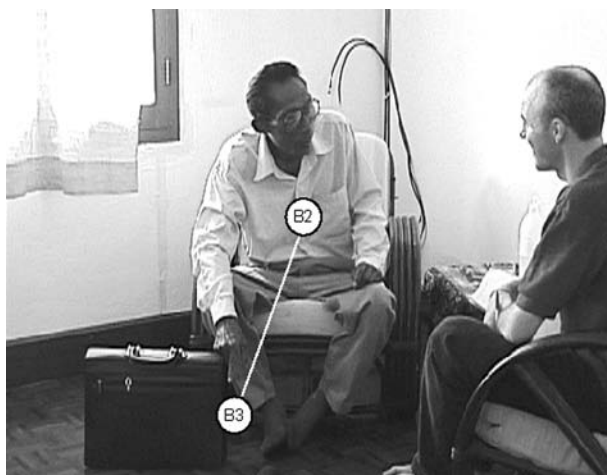


Figure 7.23 Diagram of two brothers, second-born and third-born.

was introduced as a discourse referent earlier in the sequence (see the lower white circle in Figure 7.20). The previous gestural reference to the first brother was a pointing gesture to the very low space directly in front of the speaker (Figure 7.19d). This space was set up at a time when the spatial layout of the diagram was motivated by a simple mapping of the father–son relation onto a vertical line, with father above and son below (Figure 7.20). Since then, however, two things have happened. First, the chunk of space directly in front was on one occasion used to refer to the speaker himself as the second brother (Figure 7.21c). Second, every other gestural reference to the speaker himself has been made by the speaker pointing to his own body, thus establishing his own body as the physical point of reference to himself as second brother. Thus, the speaker now physically represents himself in the diagram (i.e. he has brought himself-in-the-interaction and himself-in-the-diagram into alignment, in contrast to their earlier spatial separation in Figures 7.21b and 7.21c). Thus, the diagram at this moment looks like Figure 7.23.

The semiotic motivation of the diagram is now to depict relations between ranked siblings, with the result that Mr Naak needs the first brother (as B1) to be higher than himself, B2 (but not vertically above; cf. Figure 7.18). He is currently holding a right-hand index-finger pointing gesture at the chunk of space to his lower right (representing B3), and this gesture is held fast throughout his next series of gestures. (Compare the function of gesture holds as an enchronic glue between moves in sequence, in Chapter 5, above.)

Mr Naak's problem is to create a new chunk of space referring to his older brother, B1, since the space he used before – in front and below – is no longer appropriate or available, having been deactivated due to the placement of a



Figure 7.24 (a) ‘Going from the [-] first one . . .’; (b) ‘. . . that’s the first [qaaj4] . . .’; (c) ‘. . . [I] . . .’; (d) ‘. . . call him the first [qaaj4]’.

new point of reference in the same space (Figure 7.21c). His solution is to annex space in the left extreme periphery of his gesture space (appropriately representing B1’s status as higher than but not directly above the other brothers; cf. Chapter 6, above). The four deictic gesture strokes depicted in Figures 7.24a–7.24d refer, respectively, to B1, B1, B2 (i.e. the speaker himself), and then B1.

- (19) *tòd1 caak5 phuø thii1 [-] nùng1 – mèn1 [qaaj4] phuø*
 connect from MC.HUM ORD one be qaaj4 MC.HUM
thii1 nùng1 – [khòj5] qeen4 [qaaj4] phuø thii1 nùng1
 ORD one 1SG.P call qaaj4 MC.HUM ORD one
 ‘Going from the[-]_{Fig. 7.24a} first one – that’s the first [qaaj4]_{Fig. 7.24b} –
 [I]_{Fig. 7.24c} call him the first [qaaj4]_{Fig. 7.24d} . . .’

Figures 7.24a–7.24d illustrate a hold-and-work-with-free-hand strategy (as in Example 1, above), in which the moving hand represents something being asserted in the discourse (relating directly to what is being said), while the non-moving hand forms a held deictic gesture anchored on a referent that had just been active and will be returned to, but that during this section of the discourse is backgrounded (cf. Chapter 5).

Having made this digression about his first brother, B1, establishing a whole new position for him in the diagram, Mr Naak turns his attention now back to B3, the referent of the right-hand deictic gesture he set up in Figure 7.22a and has held in place for over twenty seconds throughout the sequence illustrated in Figures 7.22c–7.24d. By the time he begins to speak, his gaze is already back on B3’s chunk of space, indicating that he is now foregrounding it as a referent in the utterance currently underway. The deictic gestures, pointing to B3 and the speaker (as B2) respectively, are shown in Figures 7.25a and 7.25b.

- (20) *[tòd1] caak5 khòj5 long2 paj3 qiik5 – [qeen4] qaaj4 –*
 connect from 1SG.P descend go more call qaaj4
qeen4 khòj5 pèn3 qaaj4 phuø thii1 sòdng3
 call 1SG.P be qaaj4 MC.HUM ORD two
 ‘(The one) [connecting]_{Fig. 7.25a} further down from me – [calls]_{Fig. 7.25b}
 (me) qaaj4 – calls me the second qaaj4.’



Figure 7.25 (a) ‘(The one) [connecting] further down from me . . .’; (b) ‘. . . [calls] (me) *qaaj4* – calls me the second *qaaj4*’; (c) ‘[Connecting] on from the thir – the second going down is the third one . . .’; (d) ‘. . . [that] one’.

Finally, Mr Naak introduces yet another younger brother in order to reiterate the point that a brother who is called *nòòng4* ‘younger sibling’ by his elder sibling will still be called *qaaj4* ‘older brother’ by a younger sibling. This fourth brother is placed even farther to the right in Mr Naak’s developing diagram. In order to add this additional node, Mr Naak uses a place-mark technique to maintain a distinction between the existing B3 node (held in place throughout Figures 7.22c–7.24d) and a new B4 node, further out to the right. He first points with his right hand to the B3 node (Figure 7.25c) and then points momentarily with his left hand to his own right-hand gesture (Figure 7.25d), as if to hold that chunk of space down, thus giving the right hand freedom to move across and create a new and distinct chunk of space with a new and distinct referent.

- (21) [tòd1] *caak5 phuø thii1 sa – sòòng3*
 connect from MC.HUM ORD thr – two
long2 paj3 pên3 phuø thii1 saam3 – [phuø
 descend go is MC.HUM ORD three MC.HUM
nan4] –
 DEM.NONPROX

‘[Connecting]Fig. 7.25c on from the thir – the second going down is the third one – [that one]Fig. 7.25d . . .’

He now points to a new chunk of space in the low right periphery, creating a new node in the diagram for B4, then pointing once again to B3, as he explains the terms these two brothers will use for each other. Figures 7.26a and 7.26b show the preparation and stroke for the gesture referring to the fourth brother. Figures 7.26c and 7.26d show the same for the third brother. The speaker’s gaze follows the deictic gestures.

- (22) [laø] [ñang2] *mii2 nòòng4 qiik5 phuø nùng1 qiik5 baat5-niø*
 PRF still there.is nòòng4 more MC.HUM one more THZR



Figure 7.26 (a) ‘... [and] there is ...’; (b) ‘... [still] another *nòng4* ...’; (c) ‘... (who) must ...’; (d) ‘... [call] this person *qaaj4*, the third one’.

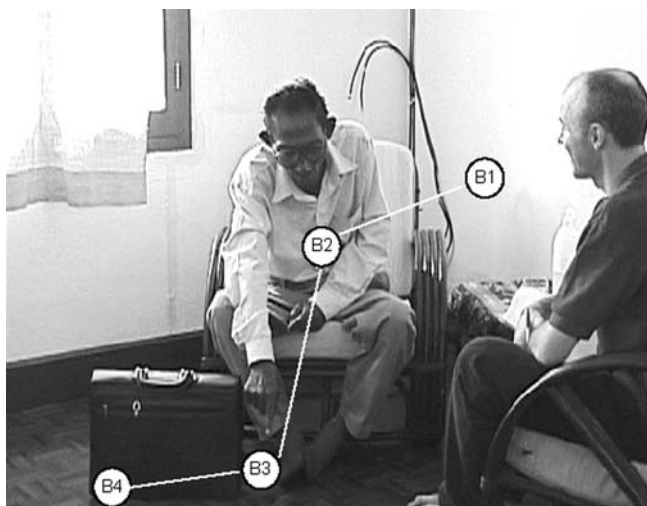


Figure 7.27 Diagram of four brothers, ranked in age.

– [tòòng4] [qeen4] phuø-nii4 pên3 qaaj4 phuø thii1
 OBLIG call MC.HUM-DEM be qaaj4 MC.HUM ORD
saam3
 three

‘... [and]_{Fig. 7.26a} there is [still]_{Fig. 7.26b} another *nòng4* – (who)
 [must]_{Fig. 7.26c} [call]_{Fig. 7.26d} this person *qaaj4*, the third one’.

The preparation phases for these two pointing gestures (shown in Figures 7.26a and 7.26c) are exaggerated (i.e. they are raised vertically higher than might be expected), presumably helping to accentuate the distinctness in the diagram between these two rather proximate chunks of space, referring to the third and fourth brothers.

The final state of the diagram is made explicit in Figure 7.27 (cf. Figure 7.18).

7.4 Discussion of example 2

As in Example 1, and in the other kinship examples discussed in Chapter 6, above, the speaker creates a complex virtual diagram in gesture space, using pointing gestures to activate chunks of space and assign reference to them. In this case, the speaker exploits the vertical dimension in metaphorically mapping relational structure onto gesture space. Though virtual, the diagram displays spatial cohesion and temporal persistence. Points on it remain active while not being gestured to, and become targets for later deictic gestures.

Also as in Example 1, the speaker in this example finds himself having to restructure or edit his diagram as it develops. This is due firstly to physical constraints of the gesture space, emergent only in the diagram's enchronic development, and secondly to changing metaphorical mapping of the diagram's spatial structure to its intended meaning. Mr Naak uses the techniques of hold-and-work-with-free-hand (Figures 7.24a–7.25a; cf. also Figures 7.6b–7.9b in Example 1) and move-body-into-new-space (compare Figures 7.19a and 7.26b; the distinction in this example is more subtle than in Example 1 – cf. Figures 7.9a and 7.9b). Mr Naak also employs a further editing technique, namely reassignment-of-old-chunk-to-new-chunk. This refers to the case in which a chunk of space previously employed to stand for a certain referent is no longer usable for some reason (here the space is too crowded and the semiotic/metaphoric motivation for the placement of the chunk in question has changed), and the speaker simply creates another chunk with the same reference elsewhere in the gesture space (Figures 7.19d, 7.24a, and 7.24b). The referent of the gesture is not new but is currently semi-active, having already been established in the discourse.

7.5 A note on eye gaze

Throughout this chapter, I have repeatedly mentioned the involvement of speakers' eye gaze in the construction and maintenance of gesture diagrams. High sensitivity to the direction of others' gaze is a primordial feature of human perception and social cognition (Baron-Cohen 1995: 38ff.). Since gaze provides a directional vector, it can constitute a deictic gesture in itself (see Chapter 3, above). And it can be an especially powerful means to compound the attention-directing and focusing function of manual deictic gestures. Many mentions of gaze in the above sequences involve this kind of compounding of deictic function, in which two deictic vectors converge, one projected by the hand and one by the eyes.

The function of deictic gaze in constructing gesture diagrams may not only be to strengthen the referent-creating indexical power of a deictic hand gesture, but may also be to allow more accurate establishment of referential chunks in three-dimensional space. If a speaker simply hand-points into nearby space without looking, an addressee can imagine a vector being projected, but cannot know from form alone how far along that vector the intended referent is located. (When a referent is located at greater distances, people can employ

angle of the pointing arm for signifying distance; i.e. higher angle of a pointing arm signals that a referent is further away.) However, if one simultaneously hand-points and casts one's gaze in a given direction, one exploits two vectors originating at different points on a vertical plane. The respective angles of these two vectors can be varied such that they will converge at different distances from the body. This allows an onlooker to determine how far away from a speaker a referential chunk of space – which is not physically differentiated from surrounding space and whose location, therefore, cannot be determined by direct perceptual evidence – is intended to be. An example of such manipulation of the angle of gaze can be seen in the later part of the first sequence described above, in which Mr Phouthong maintains four separate referential chunks of space on a flat plane in front of his body. His use of different angles of gaze helps to provide the extra accuracy required for keeping these chunks distinct in what is becoming a crowded space; compare Mr Phouthong's downward gaze at his children in Figures 7.15a and 7.15b with his outward gaze to his grandchildren in Figure 7.3c and in Figures 7.14a and 7.14b.

A further use of gaze in these sequences is to signal and direct attention to features of the diagram as relevant. In (20), above, Mr Naak resumes discussion of a referent, B3, that has been backgrounded momentarily, but is visually held in play by a held index-finger deictic gesture. As he brings this referent back into the foreground of discourse with his speech, Mr Naak simultaneously directs his gaze to the relevant node on the diagram (illustrated in Figure 7.25a), signalling his renewed attention to it in the exposition.

7.6 Concluding remark

Movements of the hands can be used to create structured arrangements of reference-activated chunks of space in otherwise undifferentiated gesture space. Such virtual diagrams, whose existence is sustained by the minds of interlocutors, have temporal persistence and spatial cohesion, and accordingly they can be manipulated (e.g. shifted, resized, rotated) by movements of the hands. Further attesting to their structural robustness, such diagrams may need to be overtly edited. Features of a gesture-created diagrammatic configuration can be altered while persistence and general cohesion of the diagram is maintained. The need for adjustment may be due to change in what is intended (i.e. errors, change of mind), unforeseen constraints of the gesture space, or changing semiotic motivation for the layout of the diagram (e.g. due to competing metaphors).

The latter two motivations were observed in the data presented above. Four techniques for editing were used: (i) hold-and-drag; (ii) hold-and-work-with-free-hand; (iii) reassignment-of-old-chunk-to-new-chunk; and (iv) move-body-into-new-space. There are likely to be further means for editing or otherwise manipulating diagrams and models produced with co-speech gesture in sequences of composite moves. This is a matter for further research.

8 CONCLUSION

In battle there are only the normal and extraordinary forces, but their combinations are limitless.

Sun Tzu

8.1 The anatomy of meaning: summary remarks

In solving the ever-present puzzle of figuring out what others are trying to say, our evidence comes in chunks: composite utterances built from multiple signs of multiple types. These composites are produced by people on trajectories of collaborative social activity. As communicative behaviours, they are strategic, context-embedded efforts to make social goals recognizable. The essential claim of this book is that if we are to understand how people interpret such efforts, our primary unit of analysis must be the utterance or move, the single increment in a sequence of social interaction. Component signs will only make sense in terms of how they contribute to the function of the move as a whole.

This book has focused on moves built from speech-with-gesture as a sample domain for exploring the anatomy of meaning. But the analytic requirement to think in terms of composite utterances is not unique to speech-with-gesture. Because all utterances are composite in kind, our findings on speech-with-gesture should help us to understand meaning more generally. This is because research on the comprehension of speech-with-gesture is a sub-field of a more general pursuit: to learn how it is that interpreters understand token contributions to situated sequences of social interaction (cf. Schegloff 1968, Goffman 1981). Chapter 1 asked how multiple signs are brought together in unified interpretations, framing the issue in terms of semiotic function of a composite's distinct components (see Figure 1.7). A broad distinction was made between conventional meaning and non-conventional meaning, where these two may be joined by indexical mechanisms of various kinds. Think of a painting like Bouguereau's *Wave* (Figure 1.3): a title (words, conventional) is taken to belong with an image (an arrangement of paint, non-conventional) via indexical links (spatial co-placement on a gallery wall, putative source in

a single creator and single act of creation). The speech-with-gesture composites discussed in this book can be analysed in the same way. When a man says *Make it steep like this* with eye gaze fixed on his arm held at an angle (see Figure 1.4), the conventional signs of his speech are joined to the non-conventional sign of his arm gesture by means of indexical devices including temporal co-placement, source in a single producer, eye gaze, and the symbolic indexical expression *like this*. Or when a man says *Make it fluted at the mouth* with eye gaze fixed on his symmetrical ‘fluted’ hands (see Figure 1.5), conventional word and non-conventional hand are joined indexically by temporal co-placement, source in a single producer, and eye gaze. In these illustrative gesture cases (see Part II of this book), hand movements constitute the non-conventional ‘image’ component of the utterance. By contrast, in deictic gesture or pointing cases (see Part I), hand movement provides the indexical link between words and some image or thing in the world, such as buffaloes walking by (see Figure 2.9), or diagrams in ink or mid-air (see Chapters 6 and 7).

This semiotic framework permits systematic comparison of speech-with-gesture moves to other species of composite utterance. An important case is sign language of the Deaf. There is considerable controversy as to how, if at all, gesture and sign language are to be compared (cf. Emmorey and Reilly 1995). The present account makes it clear that the visible components of a sign language utterance cannot be compared directly to the visible hand movements that accompany speech, nor to mere speech alone (with visible hand movements subtracted), but may only be properly compared to the entire speech-with-gesture composite (cf. Liddell 2003). The unit of comparison must be the move. By the analysis advanced here, different components of a sign language move will have different semiotic functions, in the sense just discussed: conventional signs with non-conventional signs, linked indexically. Take the example of sign language ‘classifier constructions’ or ‘depicting verbs’ (Liddell 2003: 261ff.). In a typical construction of this kind, a single articulator (the hand) will be the vehicle for both a conventional sign component (a conventionalized hand shape such as the ASL ‘vehicle classifier’) and a non-conventional sign component (some path of movement, often relative to a contextually established set of token spatial referents), where linking indexical mechanisms such as spatio-temporal co-placement and source in a single creator are maximized through instantiation in a single sign vehicle, i.e. one and the same hand.

Another domain in which a general composite utterance analysis should fit is in linguistic research on syntax. Syntactic constructions, too, are made up of multiple signs, where these are mostly the conventional signs of morphemes and constructions (though note of course that many grammatical morphemes are symbolic indexicals). An increasingly popular view of syntax

takes lexical items (words, morphemes) and grammatical configurations (constructions) to be instances of the same thing: linguistic signs (Langacker 1987, Goldberg 1995, Croft 2001). From this viewpoint, interpretation of speech-only utterances is just as for speech-with-gesture. It means dealing with multiple, simultaneously occurring signs (e.g. *That guy* may be both noun phrase and sentential subject), and looking to determine an overall target meaning for the communicative move that these signs are converging to signify. A difference is that while semantic relations within grammatical structures are often narrowly determined by conventions like word order, speech-with-gesture composites appear to involve mere co-occurrence of signs, with no further formal instruction for interpreters as to how their meanings are to be unified. Because of this extreme under-determination of semiotic relation between, say, a gesture and its accompanying speech, many researchers conclude that there are no systematic combinatorics in speech-with-gesture. But speech-with-gesture composites are merely a limiting case in the range of ways that signs combine: all an interpreter knows is that these signs are to be taken together, but there may be no conventionally coded constraints on **how**. Such under-determination is not unique to gesture. In language, too, we find minimal interpretive constraints on syntactic combinations within the clause, as documented for example by Gil (2005) for extreme isolating grammar found in some spoken languages. And beyond the clause level, such under-determined relations are the standard fabric of textual cohesion (Halliday and Hasan 1976).

In sum, to understand the process of interpreting any type of composite utterance, we do not begin with components like noun, rising intonation, or pointing gesture. We begin instead with the notion of a whole utterance, a complete unit of social action which always has multiple components, which is always embedded in a sequential context (simultaneously an effect of something prior and a cause of something next), and whose interpretation always draws on both conventional and non-conventional signs, joined indexically as wholes.

8.2 Semiotic unification – towards a rational, heuristic-based model

This book is a case study in the general problem of semiotic unification: how is it that an interpreter takes multiple signs to stand together for a coherent whole meaning? To treat the problem in the most general terms possible, we can consider utterance comprehension as a species of decision-making. The interpreter's task is to decide what someone is trying to say (or what someone is trying to do by what they are saying). Assume that interpreters in social

interaction follow the same three-step strategy they follow in other decision-making domains (Gigerenzer *et al.* 1999):

- 1 **lock-on** to a problem to be solved (i.e. identify a target decision to be made and instigate a search);
- 2 **narrow the search** as far as possible;
- 3 **lock-off** by making a decision appropriate for current purposes, thereby stopping the search.

In the domain of interpreting others' contributions to conversation (typically, composites of speech and visible behaviour), these phases in a process of rational decision-making can be thought of as follows:

- 1 lock-on by recognizing that a swatch of behaviour has a communicative intention, and instigate a search for the informative intention behind it;
- 2 use sign filtration to constrain the search for relevant signs within that behaviour at lowest cost within a search space;
- 3 lock-off by making a currently appropriate interpretation of the utterance which constitutes a satisfactory stop to the search.

An adequate account of semiotic unification in the interpretation of addressed communicative utterances will have to map out how the three phases of rational interpretation are navigated in the enchronic environment of conversation. The following paragraphs offer some initial thoughts in this direction, based on considerations raised in Chapter 1, and explored in the empirical studies in Chapters 2–7.

8.2.1 *On-switch*

What causes an interpreter to lock on to a problem of communicative interpretation in the first place? The answer is any sign of communicative intention, i.e. any indication that a person is carrying out a controlled, addressed behaviour in order that the behaviour be taken by another person to be a sign of saying something (Grice 1957). A simple heuristic for recognizing communicative intention is to identify a behaviour whose very existence is already dedicated to that function – for example, any conventional sign such as the words and constructions of language. Other heuristics include simple cues associated with the everyday use of such signs. For example, if someone is looking straight at your face when they carry out some action, it is more likely than not that they are inviting you to attend to that action (for example they may be doing the action in order to demonstrate something to you; Csibra and Gergely 2006, Gergely and Csibra 2006).

Further heuristics for recognizing when to lock on to a behaviour and begin a search for meaning include general abductive mechanisms of rational interpretation. For example, if a person's action is done in an unusual or otherwise marked manner, or in a way that cannot be explained for obvious practical reasons, then an interpreter may conclude that this person must be doing this marked action for some special purpose, and lock on to this as a puzzle which deserves to be solved (Grice 1975). Prelinguistic infants do this when seeing adults carry out actions in unconventional ways such as turning on a light using the head rather than the hands (Gergely *et al.* 2002). Or suppose I'm telling you about someone I just met and I say, suddenly speaking at noticeably increased volume: HE SPEAKS REALLY LOUD. Your sign filter will take the increased volume to mean something, and by regarding the words together with the non-conventional sign of suddenly increased volume (indexically cemented by co-occurrence in a single sign vehicle: the speech stream), the composite utterance will readily be taken to both describe and illustrate how he speaks.¹

The types of composite utterance described in this book possess most or all of the features which these lock-on heuristics would exploit. They are thus straightforwardly taken by interpreters to be communicative problems in need of interpretive solutions. Of course, these utterances are not mere passing puzzles of interest. Interlocutors are highly motivated to deal adequately with others' moves due to the social consequences inherent in each increment to a sequence of social interaction (Schegloff 1968, Heritage and Atkinson 1984, Enfield 2006), and the morally-grounded joint commitment which social interaction entails (Clark 2006).

8.2.2 *Search: sign filtration*

Once an interpreter has locked on to a person's actions as communicatively intended, they need to know which are the signs that they should attend to, and which are not, in order to assess a signer's (putative) informative intentions – i.e. the content of what they're trying to say.² As Kendon points out (e.g. 1986, 2004), interpreters display finely tuned differentiation of attention when faced with speech and visible behaviour together. Again, conventional signs like linguistic items are straightforwardly recognized. Their very *raison d'être* is to convey communicative and informative intentions

¹ A variant would be HE SPEAKS **LIKE THIS**, using the symbolic indexical 'like this' to explicitly send you on a search for how it is that he speaks – i.e. supplying only an illustration, not a description (Engle 1998, Clark 1996).

² It does not matter whether a speaker actually has the psychological state corresponding to what is putatively conveyed by the signs. For instance, applause may be insincere, but it always **claims** to praise.

in the service of social action. (The same goes for pointing gestures, since as symbolic indexicals they incorporate conventional meaning.)

Now when a person produces conventional signs, they also produce a stream of other available information in the form of non-conventional signs. Consider the hand gestures described in Part II of this book. These, too, provide no significant sign filtration problem for interpreters, since they are gazed at, pointed to, spoken about, and positioned squarely in the attentional field of their intended addressees (cf. Clark 1996, Goodwin 2000a). In addition, these hand movements are straightforwardly connected to what the speaker is saying, in timing and in speakers' dynamic investment in the communicative activity (Levy and Fowler 2000). As such they are more likely to be taken as signs of a speaker's informative intention than, say, a random scratch of the head. Also, it seems clear that practices of pointing, tracing, diagramming, and modelling with the hands as exemplified in this book are themselves directly recognizable as conventional communicative practices (Kendon 2004, Wilkins 2006).

8.2.3 *Off-switch*

What determines that our ongoing interpretation of an ensemble of signs is sufficient for current purposes, such that no further interpretation is required?³ This is the sixty-four-thousand-dollar question. Few have given it serious attention (a notable exception being Sperber and Wilson 1995 [1986]). One thing for sure is that interpreters do not routinely make an exhaustive analysis of all signs presented in a given move (Sanford and Sturt 2002). Lock-off cannot be contingent upon a complete reading of all signs, conventional or otherwise, in a composite utterance. If it were, we would get stuck hanging on every potentially meaningful element of someone's action, unable to move forward.

If an interpreter is applying fast and frugal heuristics (Gigerenzer *et al.* 1999), the point is to lock-off on an interpretation which is good enough for current purposes. The objective is an optimal trade-off between minimizing effort and maximizing yield. Perceived communicative needs of the context will determine what is an optimal degree of attention paid to available signs and their conceivable meanings. Now if interpretations can be functionally adequate on the basis of less-than-exhaustive analyses of the available signs, this means that there are components of utterances that are dispensable. Can any generalizations be made as to what kinds of utterance components are more dispensable than others? Research is needed here.

³ Lock-off by interpreters is a correlate of McNeill's 'stop order' for producers (McNeill 2005: 18).

To understand how lock-off comes about, it is worth being mindful of the fact that interpreters do not work (purely) from the ground up. From the start, an interpreter's search is for something which **gives rise to** a producer's signs, that is, the intention which those signs should make recognizable. Upon recognition that a signer is saying something (i.e. recognition that there is a communicative intention) an interpreter searches for an informative intention. This will not be found by merely decoding and summing up the available signs, as if they were a carbon copy of the speaker's token meaning. That meaning can only be **suggested by** the signs. How we get from sheer form to speaker-meaning remains unknown, but the mechanism is bound to be simpler than it appears. An over-arching principle for lock-off in discovering what others mean (in the spirit of Sperber and Wilson 1995 [1986]) might be as follows:⁴ when you see a set of signs, and see them **as** signs, take them to stand for the first object that they could simultaneously stand for.

8.2.4 *A heuristics-based approach*

With an eye to further directions in research on composite utterances, I have hinted at some elements of a heuristics-based approach to analysing how moves as social actions are understood on the basis of formal patterns of behaviour in interaction. In Chapter 1, I suggested a preliminary list of interpretive triggers and heuristics, including a convention heuristic, an orientation heuristic, a contextual association heuristic, a unified utterance–meaning heuristic, and an agency heuristic. We can also expect to rely heavily on a relevance heuristic, inherent in an enchronic model of meaning and its emphasis not just on moves but on relations between moves. A central task for subsequent research is to discover, describe, and test the full set of heuristics which make tractable the mind-boggling yet seemingly effortless task of figuring out what others are trying to say.

8.3 **Concluding remark: speech, gesture, and meaning**

Research on speech-with-gesture yields ample motivation to question the standard focus in mainstream linguistics on competence and static representations of meaning (as opposed to performance and dynamic processes of meaning; see McNeill 2005: 64ff., Wilkins 2006: 140–141). There is a need for due attention to meaning at a context-situated token level (a stance preferred by many functionalist linguists, linguistic anthropologists, conversation

⁴ With thanks to Paul Kockelman for this pithy phrasing of a thought that once took me an hour to express. Note that 'object' here is meant in the technical sense of Peirce (1955); i.e. whatever a sign stands for (not necessarily a physical object).

analysts, and some gesture researchers), rather than privileging the analysis of abstract types (preferred by a Saussurean mainstream). Speech-with-gesture composites quickly make this need apparent, because they force us to examine singularities, i.e. semiotic structures that are tokens but not tokens-of-types. These singularities include non-conventional gestures as utterance components, as well as the overall utterances themselves, each a unique combination of signs. This is why, for instance, Kendon writes of speech-with-gesture composites that 'it is only by studying them as they appear within situations of interaction that we can understand how they serve in communication' (2004: 47–48). Here is the key point: what Kendon writes is already true of speech whether it is accompanied by gesture or not (e.g. Hanks 1990, 1996 among many others). It is just that these lessons are not taught in mainstream linguistics. Speech-with-gesture teaches us to address meaning at the token-level, something we should be doing anyway.

The point is reminiscent of Theodore Roszak's (1977) argument that the changes in social practice necessary to improve our health and well-being locally are the same changes necessary to avert environmental disaster globally. Or as Roszak put it: 'The needs of the person are the needs of the planet.' In the same spirit, the analytic stance that speech-with-gesture demands – i.e. to treat moves as dynamic, motivated, concrete, and context-bound – is the same stance we need for the proper treatment of communicative moves generally, including the subject matter of linguistics, anthropology, and other branches of semiotics. In other words: the needs of research on gesture are the needs of research on meaning.

Bibliography

- Anderson, S. and Keenan, E. 1985, Deixis, in T. Shopen (ed.), *Language typology and syntactic description*, Vol. III: *Grammatical categories and the lexicon*, pp. 259–307. Cambridge: Cambridge University Press.
- Armstrong, D.F., Stokoe, W.C., and Wilcox, S.E. 1995, *Gesture and the nature of language*. Cambridge: Cambridge University Press.
- Atlas, J.D. 2005, *Logic, meaning, and conversation: Semantical underdeterminacy, implicature, and their interface*. Oxford: Oxford University Press.
- Austin, J.L. 1962, *How to do things with words*. Cambridge, MA: Harvard University Press.
- Baddeley, A.D. 1986, *Working memory*. Oxford: Clarendon Press.
- Baron-Cohen, S. 1995, *Mindblindness: An essay on autism and theory of mind*. Cambridge, MA/London: MIT Press.
- Barr, D.J. and Keysar, B. 2004, Making sense of how we make sense: The paradox of egocentrism in language use, in H. Colston and A. Katz (eds.), *Figurative language processing: Social and cultural influences*, pp. 21–41. Mahwah: Erlbaum.
- Bates, E., Camaioni, L., and Volterra, V. 1975, The acquisition of performatives prior to speech, *Merril-Palmer Quarterly* 21: 205–224.
- Bates, E., O'Connell, B., and Shore, C.M. 1987, Language and communication in infancy, in J.D. Osofsky (ed.), *Handbook of infant competence (second edition)*, pp. 149–203. New York: Wiley and Sons.
- Battison, R. 1978, *Lexical borrowing in American sign language*. Silver Spring, MD: Linstok Press.
- Bavelas, J.B. 1994, Gestures as parts of speech: Methodological implications, *Research on Language and Social Interaction* 27: 201–221.
- Bavelas, J.B., Chovil, N., Lawrie, D.A., and Wade, A. 1992, Interactive gestures, *Discourse Processes* 15: 469–489.
- Beattie G. 2003, *Visible thought: The new psychology of body language*. London and New York: Routledge.
- Binkofski, F. and Buccino, G. 2004, Motor functions of the Broca's region, *Brain and Language* 89(2): 362–369.
- Birdwhistell, R.L. 1970, *Kinesics and context: Essays on body motion communication*. Philadelphia: University of Pennsylvania Press.
- Bloch, M. and Sperber, D. 2002, Kinship and evolved psychological dispositions: The mother's brother controversy reconsidered, *Current Anthropology* 43: 723–748.
- Bonda, E., Petrides, M., and Ostry, D. 1996, Specific involvement of human parietal systems and the amygdala in the perception of biological motion, *Journal of Neuroscience* 16: 3737–3744.

- Brandom, R. B. 1979, Freedom and constraint by norms, *American Philosophical Quarterly* 16(3): 187–196.
- Brown, R. 1958, *Words and things*. Glencoe: The Free Press.
- Bouquet, M. 1996, Family trees and their affinities: The visual imperative of the genealogical diagram, *Journal of the Royal Anthropological Institute* 2: 43–66.
- Bourdieu, P. 1977, *Outline of a theory of practice*. Cambridge: Cambridge University Press.
- Bowerman, M. and Levinson, S. C. (eds.) 2001, *Language acquisition and conceptual development*. Cambridge: Cambridge University Press.
- Brookes, H. 2004, A repertoire of South African quotable gestures, *Journal of Linguistic Anthropology* 14(2): 186–224.
- Brown, P. and Levinson, S. C. 1987, *Politeness: Some universals in language usage*. Cambridge: Cambridge University Press.
- Bühler, K. 1982 [1934], The deictic field of language and deictic words, in R. J. Jarvella and W. Klein (eds.), *Speech, place, and action*, pp. 9–30. New York: Wiley.
- Burenhult, N. 2000, *Gestures in Jahai: A preliminary report from the field*, Typescript, University of Lund.
- Burling, R. 1993, Primate calls, human language, and nonverbal communication, *Current Anthropology* 34: 25–53.
- Butterworth, G. 2003, Pointing is the royal road to language for babies, in S. Kita (ed.), *Pointing*, pp. 9–34. Mahwah: Lawrence Erlbaum.
- Butterworth, G. and Morissette, P. 1996, Onset of pointing and the acquisition of language in infancy, *Journal of Reproductive and Infant Psychology* 14: 219–231.
- Button, G. (ed.) 1990, *Technology in working order: Studies of work, interaction, and technology*. London and New York: Routledge.
- Calbris, G. 1990, *The semiotics of French gestures*. Bloomington: Indiana University Press.
- Call, J. and Tomasello, M. (eds.) 2007, *The gestural communication of apes and monkeys*. Mahwah: Lawrence Erlbaum.
- Carnap, R. 1947, *Meaning and necessity*. Chicago: University of Chicago Press.
- Carpenter, M., Nagell, K., and Tomasello, M. 1988, Social cognition, joint attention, and communicative competence from 9 to 15 months of age, *Monographs of the Society for Research in Child Development* 63(4).
- Casad, E. H. 1988, Conventionalization of Cora locationals, in B. Rudzka-Ostyn (ed.), *Topics in cognitive linguistics*, pp. 345–378. Amsterdam and Philadelphia: John Benjamins.
- Chafe, W. 1994, *Discourse, consciousness, and time: The flow and displacement of conscious experience in speaking and writing*. Chicago: University of Chicago Press.
- Chao, Y. R. 1930, A system of tone letters, *Le maître phonétique, troisième série* 30: 24–27.
- Cheney, D. and Seyfarth, R. M. 1996, Function and intention in the calls of non-human primates, in W. G. Runciman, J. Maynard Smith and R. I. M. Dunbar (eds.), *Evolution of social behaviour patterns in primates and man*, pp. 56–76. Oxford: Oxford University Press.
- Claridge, G. F., Sorangkoun, T., and Baird, I. G. 1997, *Community fisheries in Lao PDR: A survey of techniques and issues*. Vientiane: IUCN.

- Clark, A. 2002, Towards a science of the bio-technical mind, *International Journal of Cognition and Technology* 1(1): 21–33.
- Clark, E. V. 1999, Acquisition in the course of conversation, *Studies in the Linguistic Sciences* 29(2): 1–18.
- Clark, H. H. 1992, *Arenas of language use*. Chicago: University of Chicago Press.
- 1996, *Using language*. Cambridge: Cambridge University Press.
- 2003, Pointing and placing, in S. Kita (ed.), *Pointing: Where language, culture, and cognition meet*, pp. 243–268. Mahwah: Lawrence Erlbaum.
- 2006, Social actions, social commitments, in N.J. Enfield and S.C. Levinson (eds.), *Roots of human sociality: Culture, cognition, and interaction*, pp. 126–152. London: Berg.
- Clark, H. H., Schreuder, R., and Buttrick, S. 1983, Common ground and the understanding of demonstrative reference, *Journal of Verbal Learning and Verbal Behavior* 22: 245–258.
- Cline, M. G. 1967, The perception of where a person is looking, *American Journal of Psychology* 80: 41–50.
- Colapietro, V. M. 1989, *Peirce's approach to the self: A semiotic perspective on human subjectivity*, Albany: State University of New York Press.
- Condon, W. S. and Ogston, W. D. 1967, A segmentation of behavior, *Journal of Psychiatric Research*: 221–235.
- Conklin, H. C. 1969 [1964], Ethnogenealogical method, in S.A. Tyler (ed.), *Cognitive anthropology*, pp. 93–122. New York: Holt, Rinehart and Winston.
- Cooke, J. R. 1968, *Pronominal reference in Thai, Burmese, and Vietnamese*. Berkeley: University of California Press.
- Corballis, M. C. 1999, The gestural origins of language, *American Scientist* 87: 138–144.
- Croft, W. 2001, *Radical construction grammar: Syntactic theory in typological perspective*. Oxford: Oxford University Press.
- 2003, *Typology and universals*, second edition. Cambridge: Cambridge University Press.
- Cruse, D. A. 1986, *Lexical semantics*. Cambridge: Cambridge University Press.
- Csibra G. and Gergely, G. 2006, Social learning and social cognition: The case for pedagogy, in Y. Munakata and M.H. Johnson (eds.), *Processes of change in brain and cognitive development: Attention and performance*, XXI, pp. 249–274. Oxford: Oxford University Press.
- D'Andrade, R. D. 1987, A folk model of the mind, in D. Holland and N. Quinn (eds.), *Cultural models in language and thought*, pp. 112–148. Cambridge: Cambridge University Press.
- Deacon A. B. and Wedgwood, C. H. 1934, Geometrical drawings from Malekula and other islands of the New Hebrides, *Journal of the Royal Anthropological Institute* 64: 129–175.
- Deacon, T. W. 1997, *The symbolic species: The coevolution of language and the brain*. New York: W.W. Norton.
- DeMarrais, K. B., Nelson, P. A., and Baker, J. H. 1992, Meaning in mud: Yup'ik Eskimo girls at play, *Anthropology and Education Quarterly* 23: 120–144.
- Dennett, D. 1983, Intentional systems in cognitive ethology: The “Panglossian” paradigm defended, *Behavioral and Brain Sciences* 6: 343–355.

- De Jorio A. 2000 [1832], *Gesture in Naples and gesture in classical antiquity. A translation of La Mimica Degli Antichi Investigata Nel Gestire Napoletano, gestural expression of the Ancients in the light of Neapolitan gesturing*. Bloomington: Indiana University Press.
- Diessel, H. 1999, *Demonstratives: Form, function, and grammaticalization*. Amsterdam and Philadelphia: John Benjamins.
- DiLucchio, L. (ed.) 1998/1999, *About life on the Navajo nation and among the Navajo people*, http://ourworld.compuserve.com/homepages/larry_dilucchio/fag02b.htm, accessed 2001.
- Dixon, R. M.W. 2003, Demonstratives: A cross-linguistic typology, *Studies in Language* 27: 61–112
- Donald, M. 1991, *Origins of the modern mind: Three stages in the evolution of culture and cognition*. Cambridge: Harvard University Press.
- Dubinskas, F. A. and Traweek, S. 1984, Closer to the ground: A reinterpretation of Walbiri iconography, *Journal of the Royal Anthropological Institute* 19(I): 15–30.
- Dunbar, R. I.M. 1988, *Primate social systems*. London and Sydney: Croom Helm.
- Eckert, P. and Hudson, J. 1988, *Wangka Wiru: A handbook for the Pitjantjatjara language learner*, University of South Australia.
- Edmondson, J. A. and Solnit, D. B. 1997, Introduction, in J.A. Edmondson and D.B. Solnit (eds.), *Comparative Tai: The Tai branch*, pp. 1–32. Arlington: Summer Institute of Linguistics and University of Texas at Arlington.
- Efron, D. 1972 [1941], *Gesture, race, and culture: A tentative study of some of the spatio-temporal and “linguistic” aspects of the gestural behavior of Eastern Jews and Southern Italians in New York City, living under similar as well as different environmental conditions*. The Hague/Paris: Mouton.
- Ekman, P. and Friesen, W. V. 1969, The repertoire of nonverbal behavior: Categories, origins, usage, and coding, *Semiotica* 1: 49–98.
- Emmorey, K. 2001, Space on hand: The exploitation of signing space to illustrate abstract thought, in M. Gattis (ed.), *Spatial schemas and abstract thought*, pp. 147–174. Cambridge: MIT Press.
- 2002, *Language, cognition, and the brain: Insights from sign language research*. Mahwah: Lawrence Erlbaum.
- Emmorey, K. and Casey, S. 2002, Gesture, thought, and spatial language, in K.R. Coventry and P. Olivier (eds.), *Spatial language: Cognition and computational perspectives*, pp. 87–101. Dordrecht: Kluwer.
- Emmorey, K. and Falgier, B. 1999, *Processing continuous and simultaneous reference in ASL*, Linguistic Society of America Annual Meeting, Los Angeles, January 1999.
- Emmorey, K. and Reilly, J. S. (eds.) 1995, *Language, gesture, and space*. Hillsdale: Erlbaum.
- Emmorey, K., Tversky, B., and Taylor, H. A. 2000, Using space to describe space: Perspective in speech, sign, and gesture, *Spatial Cognition and Computation* 2(3): 157–180.
- Enfield, N.J. 2000, The theory of cultural logic: How individuals combine social intelligence with semiotics to create and maintain cultural meaning, *Cultural Dynamics* 12: 35–64.

- 2006, Social consequences of common ground, in N.J. Enfield and S.C. Levinson (eds.), *Roots of human sociality: Culture, cognition, and interaction*, pp. 399–430. London: Berg.
- 2007a, *A grammar of Lao*. Berlin: Mouton de Gruyter.
- 2007b, Meanings of the unmarked: How default person reference does more than just refer, in N.J. Enfield and T. Stivers (eds.), *Person reference in interaction: Linguistic, cultural, and social perspectives*, pp. 97–120. Cambridge: Cambridge University Press.
- Enfield, N.J. and Levinson, S.C. 2006, Introduction: Human sociality as a new interdisciplinary field, in N.J. Enfield and S.C. Levinson (eds.), *Roots of human sociality: Culture, cognition, and interaction*, pp. 1–38. Oxford: Berg.
- Enfield, N.J. and Stivers, T. (eds.) 2007, *Person reference in interaction: Linguistic, cultural, and social perspectives*. Cambridge: Cambridge University Press.
- Enfield, N., Kita, S., and Ruiters, J.P. de. 2007, Primary and secondary pragmatic functions of pointing gestures, *Journal of Pragmatics* 39: 1722–1741.
- Engberg-Pedersen, E. 1994, Some simultaneous constructions in Danish sign language, in M. Brennan and G. Turner (eds.), *Word-order issues in sign language*, pp. 73–87. Durham: ISLA Publications.
- Engle, R. A. 1998, Not channels but composite signals: Speech, gesture, diagrams and other demonstrations are integrated in multimodal explanations, in M.A. Gernsbacher and S.J. Derry (eds.), *Proceedings of the twentieth annual conference of the Cognitive Science Society*, pp. 321–327. Mahwah: Lawrence Erlbaum.
- Feldman, H. 1986, *A grammar of Awtuw*. Canberra: Australian National University.
- Fillmore, C. J. 1982, Towards a descriptive framework for spatial deixis, in R.J. Jarvella and W. Klein (eds.), *Speech, place, and action*, pp. 31–59. Chichester: John Wiley.
- 1997 [1971], *Lectures on deixis*. Stanford: CSLI Publications.
- Floel, A., Ellger, T., and Breitenstein, C. 2003, Language perception activates the hand motor cortex: Implications for motor theories of speech perception, *European Journal of Neuroscience* 18: 704–708.
- Foley, W.A. and Van Valin Jr, R. D. 1984, *Functional syntax and universal grammar*. Cambridge: Cambridge University Press.
- Frege, G. 1960 [1892], On sense and reference, in P. Geach and M. Black (eds.), *Translations from the philosophical writings of Gottlob Frege*, pp. 56–78. Oxford: Blackwell.
- Frith, C. D. and Frith, U. 2007, Social cognition in humans, *Current Biology* 17: R724–R32.
- Gelb, I. J. 1952, *A study of writing: The foundations of grammatology*. London: Routledge and Kegan Paul.
- Gentner, D. and Goldin-Meadow, S. (eds.) 2003, *Language in mind: Advances in the study of language and thought*. Cambridge: MIT Press.
- Gergely, G. and Csibra, G. 2006, Sylvia's recipe: The role of imitation and pedagogy in the transmission of cultural knowledge, in N.J. Enfield and S.C. Levinson (eds.), *Roots of human sociality: Culture, cognition, and interaction*, pp. 229–255. London: Berg.
- Gergely, G., Bekkering, H., and Király, I. 2002, Developmental psychology: Rational imitation in preverbal infants, *Nature* 415: 755.

- Gibson, J.J. 1979, *The ecological approach to visual perception*. Boston: Houghton Mifflin.
- Gibson, J.J. and Pick, A.D. 1963, Perception of another person's looking behavior, *American Journal of Psychology* 76: 386–394.
- Gigerenzer, G., Todd, P.M., and The ABC Research Group 1999, *Simple heuristics that make us smart*. Oxford: Oxford University Press.
- Gil, D. 2005, Isolating-monocategorical-associational language, in H. Cohen and C. Lefebvre (eds.), *Categorization in cognitive science*, pp. 348–377. Amsterdam: Elsevier.
- Glenberg, A.M. and Langston, W.E. 1992, Comprehension of illustrated text: Pictures help to build mental models, *Journal of Memory and Language* 31(2): 129–151.
- Goddard, C. 1998, Bad arguments against semantic primitives, *Theoretical Linguistics* 24: 129–156.
- Godelier, M., Trautmann, T.R., and Tjon Sie Fat, F.E. (eds.) 1998, *Transformations of kinship*. Washington, DC: Smithsonian Institution Press.
- Goffman, E. 1959, *The presentation of the self in everyday life*. New York: Anchor Books.
- 1963, *Behavior in public places: Notes on the social organization of gatherings*. New York: The Free Press.
- 1964, The neglected situation, *American Anthropologist* 66: 133–136.
- 1981, *Forms of talk*. Philadelphia: University of Pennsylvania Press.
- Goldberg, A.E. 1995, *Constructions: A construction grammar approach to argument structure*. Chicago: University of Chicago Press.
- 2006, *Constructions at work: The nature of generalization in language*. Oxford: Oxford University Press.
- Goldin-Meadow, S. 1999, The role of gesture in communication and thinking, *Trends in Cognitive Sciences* 3: 419–429.
- 2003a, *Hearing gesture: How our hands help us think*. Cambridge: Belknap Press.
- 2003b, *The resilience of language: What gesture creation in deaf children can tell us about how all children learn language*. New York: Taylor and Francis.
- Goldin-Meadow, S. and Mylander, C. 1998, Spontaneous sign systems created by deaf children in two cultures, *Nature* 391: 279–281.
- Goldin-Meadow, S., McNeill, D., and Singleton, J. 1996, Silence is liberating: Removing the handcuffs on grammatical expression in the manual modality, *Psychological Review* 103(1): 34–55.
- Goldin-Meadow, S., Nusbaum, H., Kelly, S., and Wagner, S. 2001, Explaining math: Gesturing lightens the load, *Psychological Science* 12: 516–522.
- Goodwin, C. 1979, The interactive construction of a sentence in natural conversation, in G. Psathas, *Everyday language: Studies in ethnomethodology*. New York: Irvington.
- 1986, Gesture as a resource for the organization of mutual orientation, *Semiotica* 62: 29–49.
- 1994, Professional vision, *American Anthropologist* 96: 606–633.
- 1996, Transparent vision, in E. Ochs, E.A. Schegloff, and S.A. Thompson (eds.), *Interaction and grammar*, pp. 370–404. Cambridge: Cambridge University Press.
- 2000a, Action and embodiment within situated human interaction, *Journal of Pragmatics* 32(10): 1489–1522.

- 2000b, Gesture, aphasia, and interaction, in D. McNeill (ed.), *Language and gesture*, pp. 84–98. Cambridge: Cambridge University Press.
- 2002, Time in action, *Current Anthropology* 43: S19–S35.
- 2003, Pointing as situated practice, in S. Kita (ed.), *Pointing: Where language, culture, and cognition meet*, pp. 217–241. Mahwah: Lawrence Erlbaum.
- (ed.) 2003, *Conversation and brain damage*. Oxford: Oxford University Press.
- 2006, Human sociality as mutual orientation in a rich interactive environment: Multimodal utterances and pointing in aphasia, in N.J. Enfield and S.C. Levinson (eds.), *Roots of human sociality: Culture, cognition, and interaction*, pp. 97–125. London: Berg.
- Goodwin, C., Goodwin, M.H., and Olsher, D. 2002, Producing sense with nonsense syllables: Turn and sequence in conversations with a man with severe aphasia, in C.E. Ford, B.A. Fox, and S.A. Thompson (eds.), *The language of turn and sequence*, pp. 56–80. Oxford: Oxford University Press.
- Goody, E.N. (ed.) 1995, *Social intelligence and interaction: Expressions and implications of the social bias in human intelligence*. Cambridge: Cambridge University Press.
- Goody, J. 1968, Introduction, in J. Goody (ed.), *Literacy in traditional societies*, pp. 1–26. Cambridge: Cambridge University Press.
- 1977, *The domestication of the savage mind*. Cambridge: Cambridge University Press.
- Green, J. 2008, Multimodal complexity in Arandic sand story narratives. Manuscript, University of Melbourne
- Greenberg, J.H. (ed.) 1966, *Universals of grammar*. Cambridge: MIT Press.
- Grice, H.P. 1957, Meaning, *Philosophical review* 67: 377–388.
- 1975, Logic and conversation, in P. Cole and J.L. Morgan (eds.), *Syntax and semantics*, Vol. III: *Speech acts*, pp. 41–58. New York: Seminar Press.
- 1989, *Studies in the way of words*. Cambridge: Harvard University Press.
- Griffin, D.R. 1992, *Animal minds*. Chicago: University of Chicago Press.
- Gullberg, M. 2003, Eye movements and gestures in human face-to-face interaction, in J. Hyona, R. Radach, and H. Deubel (eds.), *The mind's eye: Cognitive and applied aspects of eye movement research*, pp. 685–703. Amsterdam: North Holland.
- 2008, Gestures and second language acquisition, in P. Robinson and N.C. Ellis (eds.), *Handbook of cognitive linguistics and second language acquisition*, pp. 276–305. London: Routledge.
- Gumperz, J.J. and Levinson, S.C. 1996, *Rethinking linguistic relativity*. Cambridge: Cambridge University Press.
- Haas, M.R. 1978 [1969], Sibling terms as used by marriage partners, in *Language, culture, and history: Essays by Mary R. Haas*. Stanford: Stanford University Press.
- Haiman, J. 1985, *Natural syntax: Iconicity and erosion*. Cambridge: Cambridge University Press.
- Halliday, M.A.K. and Hasan, R. 1976, *Cohesion in English*. London: Longman.
- Hanks, W.F. 1990, *Referential practice: Language and lived space among the Maya*. Chicago: University of Chicago Press.
- 1996, *Language and communicative practices*. Boulder: Westview Press.
- 2005, Explorations in the deictic field, *Current anthropology* 46(2): 191–220.

- Hanlon, R., Brown, J., and Gerstman, L. 1990, Enhancement of naming in nonfluent aphasia through gesture, *Brain and Language* 38: 298–314.
- Harris, R. 1986, *The origin of writing*, London: Duckworth.
- Hasan, R. 1968, *Grammatical cohesion in spoken and written English*. London: Longman.
- Haviland, J.B. 1979, How to talk to your brother-in-law in Guugu Yimidhirr, in T. Shopen (ed.), *Languages and their speakers*, pp.160–239. Philadelphia: University of Philadelphia Press.
- 1993, Anchoring, iconicity, and orientation in Guugu Yimithirr pointing gestures, *Journal of Linguistic Anthropology* 3(1): 3–45.
- 2000, Pointing, gesture spaces, and mental maps, in D. McNeill (ed.), *Language and gesture*, pp.13–46. Cambridge: Cambridge University Press.
- 2003, How to point in Zinacantan, in S. Kita (ed.), *Pointing: Where language, culture, and cognition meet*, pp.139–69. Mahwah: Erlbaum.
- Hayashi, M. 2005, Referential problems and turn construction: An exploration of an intersection between grammar and interaction, *Text* 25(4): 437–468.
- Heath, C.C. and Luff, P. 2000, *Technology in action*. Cambridge: Cambridge University Press.
- Heritage, J. 2007, Intersubjectivity and progressivity in person (and place) reference, in N.J. Enfield and T. Stivers (eds.), *Person reference in interaction: Linguistic, social, and cultural perspectives*. Cambridge: Cambridge University Press.
- Heritage J., and Atkinson, J.M. 1984, Introduction, in J.M. Atkinson and J. Heritage (eds.), *Structures of social action: Studies in conversation analysis*, pp.1–15. Cambridge: Cambridge University Press.
- Heritage, J. and Raymond, G. 2005, The terms of agreement: Indexing epistemic authority and subordination in talk-in-interaction, *Social Psychology Quarterly* 68(1): 15–38.
- Hewes, G. 1973, Primate communication and the gestural origin of language, *Current Anthropology* 14: 5–24.
- 1981, Pointing and language, in T. Myers, J. Laver, and J. Anderson (eds.), *The cognitive representation of speech*, pp.263–269. Amsterdam: North-Holland.
- 1996, A history of the study of language origins and the gestural primacy hypothesis, in A. Lock and C.R. Peters (eds.), *Handbook of human symbolic evolution*, pp.571–95. Oxford: Clarendon Press.
- Hickok, G. and Poeppel, D. 2000, Towards a functional neuroanatomy of speech perception, *Trends in Cognitive Science* 4: 131–138.
- Himmelman, N.P. 1996, Demonstratives in narrative discourse: A taxonomy of universal uses, in B. Fox (ed.), *Studies in anaphora*, pp.205–254. Amsterdam: John Benjamins.
- Hockett, C.F. 1960, The origin of speech, *Scientific American* 203: 89–96.
- 1987, *Refurbishing our foundations: Elementary linguistics from an advanced point of view*. Amsterdam and Philadelphia: John Benjamins.
- Horn, L.R. 1989, *A natural history of negation*. Chicago: Chicago University Press.
- Hutchins, E. 1995, *Cognition in the wild*. Cambridge: MIT Press.
- 2006, The distributed cognition perspective on human interaction, in N.J. Enfield and S.C. Levinson (eds.), *Roots of human sociality: Culture, cognition and interaction*, pp.375–398. Oxford: Berg.

- Hutchins, E. and Hazlehurst, B. 1995, How to invent a shared lexicon: The emergence of shared form-meaning mappings in interaction, in E. Goody (ed.), *Social intelligence and interaction: Expressions and implications of the social bias in human intelligence*, pp. 53–67. Cambridge: Cambridge University Press.
- Hutchins, E. and Palen, L. 1993, Constructing meaning from space, gesture, and speech, in L.B. Resnick, R. Saljo, C. Pontecorvo, and B. Burge (eds.), *Discourse, tools, and reasoning: Essays on situated cognition*, pp. 23–40. Berlin: Springer.
- Hutton, C.M. 1990, *Abstraction and instance: The type-token relation in linguistic theory*. Oxford: Pergamon Press.
- Iverson, J.M. and Goldin-Meadow, S. 1998, Why people gesture as they speak, *Nature* 396: 228.
- Jackendoff, R. 1983, *Semantics and cognition*. Cambridge: MIT Press.
- Jakobson, R. 1971, Shifters, verbals categories, and the Russian verb, in R. Jakobson (ed.), *Selected writings II: Word and language*, pp. 130–147. The Hague and Paris: Mouton.
- Jefferson, G. 1989, Preliminary notes on a possible metric which provides for a 'standard maximum' silences of approximately one second in conversation, in D. Roger and P. Bull (eds.), *Conversation: An interdisciplinary perspective*, pp. 166–196. Clevedon: Multilingual Matters.
- Kelly, S.D. 2001, Broadening the units of analysis in communication: Speech and nonverbal behaviours in pragmatic comprehension, *Journal of Child Language* 28: 325–349.
- Kelly, S.D., Kravitz, C., and Hopkins, M. 2004, Neural correlates of bimodal speech and gesture comprehension, *Brain and Language* 89: 253–260.
- Kelly, S.D., Barr, D., Church, R.B., and Lynch, K. 1999, Offering a hand to pragmatic understanding: The role of speech and gesture in comprehension and memory, *Journal of Memory and Language* 40: 577–592.
- Kelly, S.D., Iverson, J., Terranova, J., Niego, J., Hopkins, M., and Goldsmith, L. 2002, Putting language back in the body: Speech and gesture on three time frames, *Developmental Neuropsychology* 22: 323–349.
- Kemmerer, D. 1999, 'Near' and 'far' in language and perception, *Cognition* 73: 35–63.
- Kemp, J.H. 1984, The manipulation of personal relations: From kinship to patron-clientage, in H. ten Brummelhuis and J.H. Kemp (eds.), *Strategies and structures in Thai society*. Amsterdam: University of Amsterdam.
- Kendon, A. 1972, Some relationships between body motion and speech: An analysis of an example, in A.W. Siegman and B. Pope (eds.), *Studies in dyadic communication*, pp. 177–210. New York: Pergamon Press.
- 1977, *Studies in the behavior of social interaction*. Bloomington: Indiana University and Lisse: Peter de Ridder.
- 1978, Differential perception and attentional frame in face-to-face interaction: Two problems for investigation, *Semiotica* 24(3/4): 305–315.
- 1980, Gesticulation and speech: Two aspects of the process of utterance, in M.R. Key (ed.), *The relation between verbal and nonverbal communication*, pp. 207–227. The Hague: Mouton.
- 1986, Some reasons for studying gesture, *Semiotica* 62: 3–28.
- 1988, *Sign languages of Aboriginal Australia: Cultural, semiotic, and communicative perspectives*. Cambridge: Cambridge University Press.

- 1990, *Conducting interaction: Patterns of behavior in focused encounters*. Cambridge: Cambridge University Press
- 2004, *Gesture: Visible action as utterance*. Cambridge: Cambridge University Press.
- Kendon, A. and Versante, L. 2003, Pointing by hand in 'Neapolitan', in S. Kita (ed.), *Pointing: Where language, culture, and cognition meet*, pp. 109–137. Mahwah: Lawrence Erlbaum.
- Key, M. 1962, Gestures and responses: A preliminary study among some Indian tribes of Bolivia, *Studies in Linguistics* 16(3–4): 92–99.
- Keyes, C.F. 1975, Kin groups in a Thai-Lao community, in G.W. Skinner and A.T. Kirsch (eds.), *Change and persistence in Thai society*. Ithaca and London: Cornell University Press.
- Keysar, B., Barr, D., Balin, J.A., and Brauner, J.S. 2000. Taking perspective in conversation: The role of mutual knowledge in comprehension, *Psychological Science* 11: 32–38.
- Kirchner, R.S. 1979, Deixis in discourse: An exploratory quantitative study of the modern Dutch demonstrative adjectives, in T. Givón (ed.), *Syntax and semantics, Volume 12: Discourse and syntax*, pp. 355–375. New York: Academic Press.
- 1993, From meaning to message in two theories: Cognitive and Saussurean views of the Dutch demonstratives, in R.A. Geiger and B. Rudzka-Ostyn, *Conceptualizations and mental processing in language*, pp. 81–114. Berlin: Mouton de Gruyter.
- Kita, S. 2000, How representational gestures help speaking, in D. McNeill (ed.), *Language and gesture*, pp. 162–185. Cambridge: Cambridge University Press.
- (ed.) 2003, *Pointing: Where language, cognition, and culture meet*. Mahwah: Lawrence Erlbaum.
- Kita, S. and Dickey, L.W. (eds.) 1998, *Max Planck Institute for Psycholinguistics annual report 1998*. Nijmegen: Max Planck Institute for Psycholinguistics.
- Kita, S. and Özyürek A. 2003, What does cross-linguistic variation in semantic coordination of speech and gesture reveal?: Evidence for an interface representation of spatial thinking and speaking, *Journal of Memory and Language* 48: 16–32.
- Kita, S., Danziger, E., and Stolz, C. 2001, Cultural specificity of spatial schemas, as manifested in spontaneous gestures, in M. Gattis (ed.), *Spatial schemas and abstract thought*, pp. 115–146. Cambridge: MIT Press.
- Klima, E.S. and Bellugi, U. 1979, *The signs of language*. Cambridge: Harvard University Press.
- Kockelman, P. 2005, The semiotic stance, *Semiotica* 157(1–4): 233–304.
- 2006, Residence in the world: Affordances, instruments, actions, roles, and identities, *Semiotica* 162: 19–71.
- 2007, Agency: The relation between meaning, power, and knowledge, *Current Anthropology* 48(3): 375–401.
- Kosslyn, S.M. 1978, Imagery and internal representation, in E. Rosch and B.B. Lloyd (eds.), *Cognition and categorization*, pp. 217–257. Hillsdale: Lawrence Erlbaum.
- Lakoff, G. 1987, *Women, fire and dangerous things: What categories reveal about the mind*. Chicago: Chicago University Press.
- Lakoff, R. 1974, Remarks on *this* and *that*, *Chicago Linguistics Society* 10: 345–356.

- Lambrecht, K. 1994, *Information structure and sentence form: Topic, focus and the mental representation of discourse referents*. Cambridge: Cambridge University Press.
- Langacker, R. W. 1987, *Foundations of cognitive grammar, vol. 1: Theoretical prerequisites*. Stanford: Stanford University Press.
- 1997, The contextual basis of cognitive semantics, in J. Nuyts and E. Pedersen (eds.), *Language and conceptualization*, pp. 229–252. Cambridge: Cambridge University Press.
- Larkin, J. H. and Simon, H. A. 1987, Why a diagram is (sometimes) worth ten thousand words, *Cognitive Science* 11: 65–99.
- Layton, R. 1991, *The anthropology of art*, second edition. Cambridge: Cambridge University Press.
- Lee, P. 1996, *The Whorf theory complex: A critical reconstruction*. Amsterdam and Philadelphia: John Benjamins.
- Levelt, W. J. M. 1989, *Speaking: From Intention to Articulation*. Cambridge: MIT Press.
- Levinson, S. C. 1983, *Pragmatics*. Cambridge: Cambridge University Press.
- 1987, Minimization and conversational inference, in J. Verschueren and M. Bertelucci-Papi (eds.), *The pragmatic perspective*, pp. 61–129. Amsterdam: John Benjamins.
- 1995, H. P. Grice on location on Rossel Island, *Berkeley Linguistics Society* 25: 210–224.
- 1997, From outer to inner space: Linguistic categories and nonlinguistic thinking, in J. Nuyts and E. Pedersen (eds.), *Language and conceptualization*, pp. 13–45. Cambridge: Cambridge University Press.
- 2000, *Presumptive meanings: The theory of generalized conversational implicature*. Cambridge and London: MIT Press.
- 2001, Maxim, in A. Duranti (ed.), *Key terms in language and culture*, pp. 139–142. London: Blackwell.
- 2003, *Space in language and cognition: Explorations in cognitive diversity*. Cambridge: Cambridge University Press.
- Levy, E. T. and Fowler, C. A. 2000, The role of gestures and other graded language forms in the grounding of reference in perception, in D. McNeill (ed.), *Language and gesture*, pp. 215–234. Cambridge: Cambridge University Press.
- Lewis, D. K. 1969, *Convention: A philosophical study*. Cambridge: Harvard University Press.
- Liddell, S. K. 1995, Real, surrogate, and token space: Grammatical consequences in ASL, in K. Emmorey and J. S. Reilly (eds.), *Language, gesture, and space*, pp. 19–41. Mahwah: Lawrence Erlbaum.
- 1998, Grounded blends, gestures, and conceptual shifts, *Cognitive Linguistics* 9(3): 283–314.
- 2000, Indicating verbs and pronouns: Pointing away from agreement, in H. Lane and K. Emmorey (eds.), *The signs of language revisited: An anthology to honor Ursula Bellugi and Edward Klima*, pp. 303–320. Mahwah: Lawrence Erlbaum.
- 2002, Two types of directional gestures in American Sign Language: Unprompted and grammatically required. Plenary address, First Congress of the International Society for Gesture Studies, University of Texas at Austin.

- 2003, *Grammar, gesture, and meaning in American Sign Language*. Cambridge: Cambridge University Press.
- Liszkowski, U. 2006, Infant pointing at twelve months: Communicative goals, motives, and social-cognitive abilities, in N.J. Enfield and S.C. Levinson (eds.), *Roots of human sociality: Culture, cognition, and interaction*. London: Berg.
- Liszkowski, U., Carpenter, M., Henning, A., Striano, T., and Tomasello, M. 2004, Twelve-month-olds point to share attention and interest, *Developmental Science* 7(3): 297–307.
- Lucy, J. 1992a, *Language diversity and thought: A reformulation of the linguistic relativity hypothesis*. Cambridge: Cambridge University Press.
- 1992b, *Grammatical categories and cognition: A case study of the linguistic relativity hypothesis*. Cambridge: Cambridge University Press.
- Lyons, J. 1968, *Introduction to theoretical linguistics*. Cambridge: Cambridge University Press.
- 1972, Human language, in R.A. Hinde (ed.), *Non-verbal communication*, pp. 48–85. Cambridge: Cambridge University Press.
- 1977, *Semantics*. Cambridge: Cambridge University Press.
- Mandel, M. A. 1977, Iconic devices in American Sign Language, in L.A. Friedman (ed.), *On the other hand: New perspectives on American Sign Language*, pp. 57–107. New York: Academic Press.
- Marshack, A. 1972, *The roots of civilization: The cognitive beginnings of man's first art, symbol, and notation*. New York: McGraw-Hill.
- Marslen-Wilson, W. D., Levy E., and Tyler, L. K. 1982, Producing interpretable discourse: The establishment and maintenance of reference, in R. Jarvella and W. Klein (eds.), *Speech, place, and action*, pp. 339–378. Chichester: Wiley and Sons.
- Mathur, G. 2002, On the syntax-phonology interface in ASL: The case of hold. LSA Annual Meeting, 4 January 2002.
- McNeill, D. 1985, So you think gestures are nonverbal?, *Psychological Review* 92: 350–371.
- 1992, *Hand and mind: What gestures reveal about thought*. Chicago: University of Chicago Press.
- (ed.) 2000, *Language and gesture*. Cambridge: Cambridge University Press.
- 2003, Pointing and morality in Chicago, in S. Kita (ed.), *Pointing: Where language, culture, and cognition meet*, pp. 293–306. Mahwah: Lawrence Erlbaum.
- 2005, *Gesture and thought*, Chicago and London: Chicago University Press.
- McNeill, D., Levy, E. T., and Cassell, J. 1993, Abstract deixis, *Semiotica* 95(1–2): 5–19.
- Mead, G. H. 1934, *Mind, self, and society*. Chicago: University of Chicago Press.
- Miller, C. 1994, Simultaneous constructions in Quebec Sign Language, in M. Brennan and G. Turner (eds.), *Word order issues in sign language*, pp. 89–112. Durham: ISLA Publications.
- Miller, G. A. 1951, *Language and communication*. New York: McGraw-Hill.
- Miller, P. J., Fung, H., and Mintz, J. 1996, Self-construction through narrative practices: A Chinese and American comparison of early socialization, *Ethos* 24: 237–280.
- Miller, P. J., Wiley, A., Fung, H., and Liang, C. H. 1997, Personal storytelling as a medium of socialization in Chinese and American families, *Child Development* 68: 557–568.

- Morrell-Samuels, P. and Krauss, R.M. 1992, Word familiarity predicts temporal asynchrony of hand gestures and speech, *Journal of Experimental Psychology: Learning, Memory, and Cognition* 18: 615–622.
- Müller, C. 1998, Iconicity and gesture, in S. Santi, I. Guaïtella, C. Cavé, and G. Konopczynski (eds.), *Oralité et gestualité: Communication multimodale, interaction. (Orality and gestuality: Multimodal communication, interaction.)*, pp. 321–328. Paris: L'Harmattan.
- Munn, N.D. 1986 [1973], *Walbiri iconography: Graphic representation and cultural symbolism in a Central Australian society*. Chicago: University of Chicago Press.
- Noble, W. and Davidson, L. 1996, *Human evolution, language, and mind: A psychological and archaeological inquiry*. Cambridge: Cambridge University Press.
- Norman, D.A. 1988, *The design of everyday things*. New York: Basic Books.
- 1991, Cognitive artifacts, in J.M. Carroll (ed.), *Designing interaction: Psychology at the human-computer interface*, pp. 17–38. Cambridge: Cambridge University Press.
- Núñez, R.E. and Sweetser, E. 2006, With the future behind them: Convergent evidence from Aymara language and gesture in the crosslinguistic comparison of spatial construals of time, *Cognitive Science* 30(3): 401–450.
- Oestermeier, U. and Hesse, F.W. 2000, Verbal and visual causal arguments, *Cognition* 75: 65–104.
- Ohama, K. 1984, Development of pointing behavior in infants and mothers' responsive behavior: Longitudinal study of infants from 9 to 30 months, in M. Ogino, K. Ohama, K. Saito, S. Takei, and T. Tatsuno (eds.), *The development of verbal behavior VI*. Tokyo: Bulletin of the Faculty of Education, University of Tokyo.
- Okrent, A. 2002, A modality-free notion of gesture and how it can help us with the morpheme vs. gesture question in sign language linguistics (Or at least give us some criteria to work with), in R.P. Meier, K. Cormier, and D. Quinto-Pozos (eds.), *Modality and structure in signed and spoken languages*, pp. 175–198. Cambridge: Cambridge University Press.
- Olson, D.R. 1994, *The world on paper: The conceptual and cognitive implications of writing and reading*. Cambridge: Cambridge University Press.
- Ong, W.J. 1982, *Orality and literacy: The technologizing of the word*. London: Methuen.
- Owren, M.J. and Rendall, D. 2001, Sound on therebound: Returning form and function to the forefront in understanding nonhuman primate vocal signaling, *Evolutionary Anthropology* 10(2): 58–71.
- Özyürek, A. 2000, The influence of addressee location on spatial language and representational gestures of direction, in D. McNeill (ed.), *Language and gesture*, pp. 64–83. Cambridge: Cambridge University Press.
- Özyürek, A. and Kita, S. 2001, *Interacting with demonstratives: Encoding of joint attention as a semantic contrast in Turkish and Japanese demonstrative systems*. Istanbul and Nijmegen: Koç University and Max Planck Institute for Psycholinguistics, MS.
- Özyürek A., Willem, R.M., Kita, S., and Hagoort, P. 2007, On-line integration of semantic information from speech and gesture: Insights from event-related brain potentials, *Journal of Cognitive Neuroscience* 19: 605–616.
- Palmer, S.E. 1978, Fundamental aspects of cognitive representation, in E. Rosch and B.B. Lloyd (eds.), *Cognition and categorization*, pp. 259–303. Hillsdale: Lawrence Erlbaum.

- Parkin, R. 1997, *Kinship: An introduction to the basic concepts*. London: Blackwell.
- Parmentier, R. J. 1994, *Signs in society: Studies in semiotic anthropology*. Bloomington and Indianapolis: Indiana University Press.
- Pawley, A. and Syder, F. H. 2000, The one-clause-at-a-time hypothesis, in H. Riggensbach (ed.), *Perspectives on fluency*, pp. 163–191. Ann Arbor: University of Michigan Press.
- Peirce, C. S. 1955, *Philosophical writings of Peirce*. New York: Dover Publications.
- 1965 [1932], Speculative grammar, in C. Hartshorne and P. Weiss (eds.), *Collected papers of Charles Sanders Peirce*, vol. 2: *Elements of logic*, pp. 127–269. Cambridge: Belknap Press of Harvard University Press.
- Phillips, S. B. V. D., Goldin-Meadow, S., and Miller, P. J. 2001, Enacting stories, seeing worlds: Similarities and differences in the cross-cultural narrative development of linguistically isolated deaf children, *Human Development* 44: 311–336.
- Potter, J. M. 1976, *Thai peasant social structure*. Chicago and London: University of Chicago Press.
- Povinelli, D. J., Bering, J. M., and Giambrone, S. 2003, Chimpanzees' 'pointing': Another error of the argument by analogy?, in S. Kita (ed.), *Pointing: Where language, culture, and cognition meet*, pp. 35–68. Mahwah: Lawrence Erlbaum.
- Poyatos, F. 1983, *New perspectives in nonverbal communication: Studies in cultural anthropology, social psychology, linguistics, literature and semiotics*. Oxford: Pergamon Press.
- Preissler, M. A. and Bloom, P. 2008, Two-year-olds use artist intention to understand drawings, *Cognition* 106: 512–518.
- Quine, W. V. O. 1960, *Word and object*, Cambridge: MIT Press.
- 1971, On the inscrutability of reference, in D. D. Steinberg and L. A. Jakobovits (eds.), *Semantics: An interdisciplinary reader in philosophy, linguistics, and psychology*, pp. 142–154. Cambridge: Cambridge University Press.
- Reddy, M. J. 1979, The conduit metaphor: A case of frame conflict in our language about language, in A. Ortony (ed.), *Metaphor and thought*, pp. 284–324. Cambridge: Cambridge University Press.
- Rendall, D. and Owren, M. J. 2002, Animal vocal communication: Say what?, in M. Bekoff, C. Allen, and G. Burghardt (eds.), *The cognitive animal: Empirical and theoretical perspectives on animal cognition*, pp. 307–313. Cambridge: MIT Press.
- Richerson, P. J. and Boyd, R. 2005, *Not by genes alone: How culture transformed human evolution*. Chicago: University of Chicago Press.
- Richert, R. A. and Lillard, A. S. 2002, Children's understanding of the knowledge prerequisites of drawing and pretending, *Developmental Psychology* 38: 1004–1015.
- Rio, K. 2005, Discussions around a sand-drawing: creations of agency and society in Melanesia, *Journal of the Royal Anthropological Institute* 11: 410–423.
- Rolfe, L. H. 1996, Theoretical stages in the prehistory of grammar, in A. J. Lock and C. R. Peters (eds.), *Handbook of human symbolic evolution*, pp. 776–792. London: Wiley/Blackwell.
- Rosaldo, M. Z. 1980, *Knowledge and passion: Ilongot notions of self and social life*. Cambridge: Cambridge University Press.

- Roszak T. 1977, *Person/planet: The creative disintegration of industrial society*. New York: Anchor Books.
- Roth, W. 2000, From gesture to scientific language, *Journal of Pragmatics* 32: 1683–1714.
- Roth, W. and Lawless, D. V. 2000a, When up is down and down is up: Body orientation, proximity, and gestures as resources, *Language in Society* 31: 1–28.
- 2000b, Scientific investigations, metaphorical gestures, and the emergence of abstract scientific concepts, *Learning and Instruction* 12: 285–304.
- Ruiter, J. P. de. 2000, The production of gesture and speech, in D. McNeill (ed.), *Language and gesture*, pp. 284–311. Cambridge: Cambridge University Press.
- Ruiter, J. P. de and Wilkins, D. P. 1998, The synchronization of gesture and speech in Dutch and Arrernte (an Australian Aboriginal language): A cross-cultural comparison, *Oralité Et Gestualité: Communication Multimodale, Interaction*: 603–607.
- Ruiter, J. P. de, Mitterer, H., and Enfield, N. J. 2006, Projecting the end of a speaker's turn: A cognitive cornerstone of conversation, *Language* 82(3): 504–524.
- Ruiter, J. P. de, Rossignol, S., Vuurpijl, L., Cunningham, D. W., and Levelt, W. J. M. 2003, Slot: A research platform for investigating multimodal communication, *Behavior Research Methods, Instruments, and Computers* 35(3): 408–419.
- Sacks, H. 1984, On doing being ordinary, in J. M. Atkinson and J. Heritage (eds.), *Structures of social action: Studies in conversation analysis*, pp. 413–429. Cambridge: Cambridge University Press.
- 1992, *Lectures on conversation*. London: Blackwell.
- Sacks, H. and Schegloff, E. A. 1979, Two preferences in the organization of reference to persons in conversation and their interaction, in G. Psathas (ed.), *Everyday language: Studies in ethnomethodology*, pp. 15–21. New York: Irvington. (Reprint in N. J. Enfield and T. Stivers (eds.), *Person reference in interaction: Linguistic, social, and cultural perspectives*. Cambridge: Cambridge University Press, 2007.)
- Sacks, H., Schegloff, E. A., and Jefferson, G. 1974, A simplest systematics for the organization of turn-taking for conversation, *Language* 50(4): 696–735.
- Saitz, R. L. and Cervenka, E. J. 1972, *Handbook of gestures: Colombia and the United States*. The Hague: Mouton.
- Salomon, F. 2001, How an Andean 'writing without words' works, *Current Anthropology* 42: 1–27.
- Sandler, W. 1993, Hand in hand: The roles of the non-dominant hand in sign language phonology, *The Linguistic Review* 10: 337–390.
- 2002, From phonetics to discourse: The non-dominant hand and the grammar of sign language. Paper presented at LabPhon 8, New Haven, CT.
- Sanford, A. J. and Sturt, P. 2002, Depth of processing in language comprehension: Not noticing the evidence, *Trends in Cognitive Sciences* 6: 382–386.
- Saussure, F. 1959/1916, *Course in general linguistics*. New York: McGraw-Hill.
- Scaife, M. and Bruner, J. S. 1975, The capacity for joint attention in the infant, *Nature* 253: 265–266.
- Scaife, M. and Rogers, Y. 1996, External cognition: How do graphical representations work?, *International Journal of Human-Computer Studies* 45: 185–213.
- Schaffer, H. R. 1984, *The child's entry into a social world*. New York: Academic Press.

- Schefflen, A. E. 1976, *Human territories: How we behave in space-time*. Englewood Cliffs: Prentice-Hall. (Written with Norman Ashcraft.)
- Schegloff, E. A. 1968, Sequencing in conversational openings, *American Anthropologist* 70: 1075–1095.
- 1982, Discourse as an interactional achievement: Some uses of ‘uh huh’ and other things that come between sentences, in D. Tannen (ed.), *Analyzing discourse: Text and talk*, pp. 1–93. Washington, DC: Georgetown University Press.
- 1984, On some gestures’ relation to talk, in J.M. Atkinson and J. Heritage (eds.), *Structures of social action: Studies in conversation analysis*, pp. 266–296. Cambridge: Cambridge University Press.
- 2007a, Conveying who you are: The presentation of self, strictly speaking, in N.J. Enfield and T. Stivers (eds.), *Person reference in interaction: Linguistic, social, and cultural perspectives*. Cambridge: Cambridge University Press.
- 2007b, *Sequence organization in interaction: A primer in conversation analysis, volume 1*. Cambridge: Cambridge University Press.
- Schelling, T. C. 1960, *The strategy of conflict*. Cambridge: Harvard University Press.
- Schutz, A. 1970, *On phenomenology and social relations*. Chicago: University of Chicago Press.
- Searle, J. R. 1969, *Speech acts: An essay in the philosophy of language*. Cambridge: Cambridge University Press.
- Sherzer, J. F. 1973, Verbal and non-verbal deixis: The pointed-lip gesture among the San Blas Cuna, *Language in Society* 2: 117–131.
- Sherzer, J. 1983, *Kuna ways of speaking: An ethnographic perspective*. Austin: University of Texas Press.
- 1993, Pointed lips, thumbs up, and cheek puffs: Some emblematic gestures in social interactional and ethnographic context, in R. Queen and R. Barrett (eds.), *SALSA: Proceedings of the first annual symposium about languages and society – Austin*, pp. 196–211. Austin: University of Texas.
- Silverstein, M. 1976, Shifters, linguistic categories, and cultural description, in K. Basso and H. Selby (eds.), *Meaning in anthropology*, pp. 11–55. Albuquerque: University of New Mexico Press.
- Slama-Cazacu, T. 1976, Nonverbal components in message sequence: ‘Mixed syntax’, in W.C. McCormack and S.A. Wurm (eds.), *Language and man: Anthropological issues*, pp. 217–227. The Hague/Paris: Mouton.
- Sperber, D. and Wilson, D. 1995 [1986], *Relevance: Communication and cognition, second edition*. Oxford: Blackwell.
- Stivers, T. and Robinson, J. D. 2006, A preference for progressivity in interaction, *Language in Society*, 35(3): 367–392.
- Stivers, T., Enfield, N. J., and Levinson, S. C. 2007, Person reference in interaction, in N.J. Enfield and T. Stivers (eds.), *Person reference in interaction: Linguistic, cultural, and social perspectives*, pp. 1–20. Cambridge: Cambridge University Press.
- Stokoe, W. 1960, *Sign language structure: An outline of the visual communication systems of the American deaf* (= Studies in Linguistics, Occasional Papers 8). Silver Spring: Linstok Press.
- 1993, Comment on: Primate calls, human language, and nonverbal communication, by Robbins Burling, *Current Anthropology* 34: 42–43.

- Streeck, J. 1993, Gesture as communication 1: Its coordination with gaze and speech, *Communication Monographs* 60: 275–299.
- 1994, Gesture as communication 2: The audience as coauthor, *Research on Language and Social Interaction* 27: 239–267.
- 2002, A body and its gestures, *Gesture* 2: 19–44.
- Street, B. V. 1984, *Literacy in theory and practice*. Cambridge: Cambridge University Press.
- Talmy, L. 2006, The representation of spatial structure in spoken and signed language, in M. Hickmann and R. Stéphane (eds.), *Space in languages: Linguistic systems and cognitive categories*, pp. 207–238. Amsterdam: Benjamins.
- Taub, S. F. 2001, *Language in the body: Iconicity and metaphor in American Sign Language*. Cambridge: Cambridge University Press.
- Taub, S. F., Piñar, P., and Galvan, D. 2002, Comparing spatial information in speech/gesture and sign language. Presentation at First Congress of the International Society for Gesture Studies, Austin, Texas, June 2002.
- Thorpe, W. H. 1972, The comparison of vocal communication in animals and man, in R. A. Hinde (ed.), *Non-verbal communication*, pp. 27–48. Cambridge: Cambridge University Press.
- Tomasello, M. 1999, *The cultural origins of human cognition*. Cambridge: Harvard University Press.
- 2003, *Constructing a language: A usage-based theory of language acquisition*. Cambridge: Harvard University Press.
- 2006, Why don't apes point?, in N. J. Enfield and S. C. Levinson (eds.), *Roots of human sociality: Culture, cognition and interaction*, pp. 506–524. Oxford: Berg.
- Tomasello, M., Carpenter, M., Call, J., Behne, T., and Moll, H. 2005, Understanding and sharing intentions: The origins of cultural cognition, *Behavioral and Brain Sciences* 28: 675–691.
- van Geenhoven, V. and Warner N. (eds.) 1999, *Max Planck Institute for Psycholinguistics annual report 1999*. Nijmegen: Max Planck Institute for Psycholinguistics.
- Vermeerbergen, M., Leeson, L., and Crasborn, O. (eds.) 2007, *Simultaneity in signed languages: Form and function*. Amsterdam: John Benjamins.
- Vygotsky, L. S. 1962 [1934], *Thought and language*. Cambridge: MIT Press.
- Wagner, S., Nusbaum, H., and Goldin-Meadow, S. 2004, Probing the mental representation of gesture: Is handwaving spatial?, *Journal of Memory and Language* 50: 395–407.
- Wassmann, J. 1997, Finding the right path: The route knowledge of the Yupno of Papua New Guinea, in G. Senft (ed.), *Referring to space*, pp. 143–174. Oxford: Clarendon Press.
- Wierzbicka, A. 1972, *Semantic primitives*. Frankfurt: Athenäum.
- 1980, *Lingua mentalis*. Sydney: Academic Press.
- 1985, *Lexicography and conceptual analysis*. Ann Arbor: Karoma.
- 1996, *Semantics: Primes and universals*. New York: Oxford University Press.
- Wilkins, D. P. 1997, Alternative representations of space: Arrernte narratives in sand, in M. Beimans and J. van de Weijer (eds.), *Proceedings of the CLS opening academic year '97-'98*, pp. 133–164. Nijmegen: Center for Language Studies.

- 1999, Demonstrative questionnaire: 'THIS' and 'THAT' in comparative perspective, in D.P. Wilkins (ed.), *Manual for the 1999 field season*, pp. 1–24. Nijmegen: Language and Cognition Group, Max Planck Institute for Psycholinguistics.
- 2002, Integrating speech, gesture, and sand drawings: How external representation systems both influence and reflect spatial thinking. Invited talk delivered at interdisciplinary workshop 'Spatial Thinking in the Humanities and Sciences: From Perception to Meaning', CSLI and the Stanford Center for Innovative Learning, Stanford, CA, October.
- 2003, Why pointing with the index finger is not a universal (in sociocultural and semiotic terms), in S. Kita (ed.), *Pointing: Where language, culture, and cognition meet*, pp. 171–215. Mahwah: Erlbaum.
- 2006, Review of Adam Kendon (2004), *Gesture: Visible Action as Utterance* 6(1): 119–144.
- Wilkins, D. P. and Hill, D. 1995, When 'go' means 'come': Questioning the basicness of motion verbs, *Cognitive Linguistics* 6: 209–259.
- Wilkins, D. P., Kita, S., and Enfield, N. J. 2001, Ethnography of pointing questionnaire (Version 2), in S.C. Levinson and N.J. Enfield (eds.), *Manual for the field season, 2001*, pp. 136–141. Nijmegen: Language and Cognition Group, Max Planck Institute for Psycholinguistics.
- Wittgenstein, L. 1953, *Philosophical investigations*. Oxford: Basil Blackwell.
- Wundt, W. 1973 [1921], *The language of gestures* (translation of extract from *Volkerpsychologie: Eine Untersuchung der Entwicklungsgesetze von Sprache, Mythos und Sitte*, First Volume, Fourth Edition, First Part, Chapter 2. Stuttgart: Alfred Kroner Verlag, 1921). The Hague: Mouton.
- Zeshan, U. 2000, *Sign language in Indo-Pakistan: A description of a signed language*. Amsterdam: John Benjamins.
- Zhang, J. 1997, The nature of external representations in problem solving, *Cognitive Science* 21: 179–217.

Index

- abstract pointing, 107, 108, 152, 196, 199, 202, 209
acquisition, of language, 27
adaptors, 17
affiliational imperative, 92, 104
affordances, 108, 114, 115, 179, 184, 192
American Sign Language, 152, 222
Anderson, Stephen R., 25, 31, 42, 58
anthropology, 108, 150, 153, 227, 228
aphasia, 187
arbitrariness, 2, 12, 32, 33, 113
Arrente, people of Central Australia, 73, 88
Atkinson, Max, 10, 225
Atlas, Jay D., 3, 31
attention, 5, 6, 16, 19, 20, 29, 31, 40, 45, 54, 59, 91, 97, 98, 157, 179, 188, 190, 219, 220, 225
audience design, 54, 59, 61, 66
Austin, John L., 10, 11

baptism, 174, 196
Baron-Cohen, Simon, 84, 98, 219
Barr, Dale J., 31
Bates, Elizabeth, 11, 90, 107
Battison, Robbin M., 115
Bavelas, Janet Beavin, 1, 105, 107, 185
beats, 17
Beattie, Geoffrey, 1
Bellugi, Ursula, 115
B-hand, 135, 195, 196
Birdwhistell, Ray L., 1
Bloch, Maurice, 153
Bloom, Paul, 5
Bloomfield, Leonard, 111
Bolivia, 72
Bouguereau, William-Adolphe, 5, 6, 221
Bouquet, Mary, 150, 153, 191
Bourdieu, Pierre, 34
Bowerman, Melissa, 191
Brandom, Robert B., 12
Brandt, Willy, 3, 4, 8
Brookes, Heather, 1, 12

Brown, Penelope, 104, 106
Bühler, Karl, 13, 28, 29, 65, 90, 183
buoys, 114, 146
Burenhult, Niclas, 71
Burling, Robbins, 184
Butterworth, George, 90
Button, Graham, 178

Calbris, Geneviève, 1, 190
Call, Josep, ix
Carnap, Rudolf, 28
Carpenter, Malinda, 90
Cassell, Justine, 106, 107, 151, 184
Chafe, Wallace L., 11, 182
character perspective, 127, 128
chunk of space, 107, 151–3, 164, 180, 181, 182, 196–205, 196, 209–20, 221
Claridge, Gordon, 117, 118, 185
Clark, Eve V., 27
Clark, Herbert H., ix, 6, 11, 27, 28, 30, 31, 32, 42, 51, 58, 59, 61, 92, 95, 96, 99, 154, 181, 191, 225, 226
Cline, Marvin G., 84
cognitive artifacts, 149, 168, 179, 181, 183, 184, 188, 189, 190, 193
cognitive science, 108
collateral kin relations, 158, 160, 164, 166, 169, 183, 191
Colombia, 71
combinatorics of gesture, 17, 113, 147, 148, 223
common ground, 13, 15, 25, 27, 30, 37, 42, 43, 48, 49, 50, 51, 54, 58–61, 65, 104, 109, 110, 145, 156, 167, 181
communicative intention, 8, 16, 145, 224, 227
composite meaning, ix, 1, 25, 31, 48, 66, 67, 79, 80, 91–7, 100–5, 105–15, 108, 114, 116, 130, 134, 146, 147, 150, 165, 167, 176, 184, 186, 189, 190, 194, 214, 220, 227–8
composite utterance, 6, 9, 10, 13, 14, 15, 16, 20, 21, 31, 79, 80, 91, 100, 104, 105, 107, 109, 114, 130, 134, 146, 147, 150, 167, 176, 222, 223, 225, 226

- Condon, William S., 1
 Conklin, Harold C., 150, 154, 155, 166
 contrast, 62, 65, 91, 96, 115
 convention, 12, 13, 16, 18, 27, 28, 88, 227
 conventional meaning, 2, 6, 12–13, 16, 20, 27, 42, 59, 67, 87–9, 91, 100, 105, 107, 108, 114, 115, 150, 181, 221–6
 coordination problem, 27, 30, 31, 32, 35, 37, 61
 cousins, 157–63, 162, 167, 194–209, 195
 Croft, William, 223
 Crow-Omaha kinship system, 157
 Cruse, D. A., 2
 Csibra, Gergely, 224
- Danziger, Eve, 183, 190, 191
 Davidson, L., 184
 De Jorio, Andrea, 1
 De Ruiter, J. P., 1, 18, 90, 96, 108
 Deacon, Terrence W., 184
 decision-making, 223
 default, 42, 61, 83, 214
 deixis, 6, 19, 20, 23–110, 183, 195–220, 222
 demonstratives, 13, 20, 25–67, 68, 82, 95, 96
 diagrammatic mode of representation, 134, 144, 149, 153, 154, 163, 164, 166, 167, 174, 179, 184, 192, 220
 Diessel, Holger, 25, 29, 30, 31, 32, 42, 58
 differential attention, 145
 DiLucchio, Larry, 73, 86, 87
 disattent track, 145
 distal, distinction in demonstratives, 25, 32, 36, 42, 50, 51, 59, 60, 61, 65, 66, 82
 distributed cognition, 181, 186, 187
 Dixon, R. M. W., 25, 29, 31
 dominance condition, 115
 dominant hand, 116, 117, 140
 Donald, Merlin, 149, 154, 184
 Dubinskas, Frank A., 149
 Dunbar, Robin I. M., x
 dynamic view of meaning, 2, 10, 226, 227, 228
- Eckert, Paul, 73
 Efron, David, 1, 90, 182, 190, 192
 Einstein, 106, 151
 Ekman, Paul, 1, 12
 emblem gesture, 12, 13
 Emmorey, Karen, 115, 147, 150, 151, 154, 184, 194, 222
 enchrony, 10, 11, 81, 84, 108, 129, 130, 143, 146, 148, 179, 182, 183, 194, 209, 215, 219, 224, 227
 endophoric function of demonstratives, 29, 30, 32, 36
- Enfield, N. J., 2, 31, 32, 35, 83, 84, 90, 92, 104, 108, 109, 110, 156, 158, 225
 engagement area, 33, 34, 38, 40, 45–65, 66
 Engberg-Pedersen, Elisabeth, 115, 130
 Engle, Randi, ix, 1, 8, 11, 122, 225
 English, 13, 25, 29, 32, 61, 65, 88, 96, 106, 127, 154, 157, 159
 epigenesis, 10
 ethnogenealogy, 150, 154, 166, 192
 evanescence, 17, 149, 151, 154, 164, 191, 192
 exophoric function of demonstratives, 29, 30, 32, 36, 37
 external representations, 145, 179, 184
 eyebrows, 68, 69, 73, 76, 81, 84, 88, 89
- face, threat to, 105, 109
 Faller, Martina, 45
 Feldman, Harry, 68
 filial relations, 166, 168, 169, 183
 Fillmore, Charles, 12, 25, 28, 30
 first position, 106
 Foley, William A., 11
 Fowler, Carol A., 97, 105, 123, 124, 145, 226
 Frege, Gottlob, 13
 Freud, 7
 Friesen, Wallace, 1, 12
- gaze, 6, 16, 19, 38, 45, 54, 68, 69, 73–82, 84–6, 88, 89, 93, 97, 98, 120, 121, 124, 135, 149, 150, 151, 165, 167, 174, 175, 181, 185, 187, 188, 196, 198, 206, 208, 209, 211, 213, 216, 217, 219, 222
 gaze-switch, 78, 84
 Generalized Conversational Implicature, 15
 Gentner, Dedre, 191
 Georgian, 107
 Gergely, György, 16, 224, 225
 gestural layering, 114
 gesture hold, 124, 146
 gesture space, 19, 152, 153, 166, 169, 170, 172, 191, 194, 196, 203, 206, 209–13, 216, 219, 220
 G-hand, 198
 Gibson, James J., 84
 Gigerenzer, Gerd, 224, 226
 Gil, David, 223
 Glenberg, Arthur M., 108, 150, 179
 Goddard, Cliff, 29
 Godelier, Maurice, 153
 Goffman, Erving, 1, 10, 11, 33, 42, 52, 106, 153, 221
 Goldberg, Adele E., 223
 Goldin-Meadow, Susan, 1, 115, 150, 154, 184, 187, 191
 Gombrich, E. H., 149

- Goodwin, Charles, ix, 1, 11, 33, 58, 98, 149, 152, 154, 178, 181, 182, 184, 185, 187, 190, 226
- Goodwin, Marjorie H., 187
- Goody, Esther, 32
- Goody, Jack, 108, 149, 150, 178, 179, 184, 191, 192
- Green, Jenny, 149, 154
- Grice, H. Paul, 2, 3, 13, 15, 16, 26, 27, 28, 105, 224, 225
- growth point, 11
- Gullberg, Marianne, 1, 185
- Gumperz, John J., 191
- Haas, Mary R., 156
- habitus, 34
- Haiman, John, 182
- Halliday, M. A. K., 11, 25, 29, 36, 113, 147, 223
- Hanks, William F., xii, 2, 15, 25, 28, 29, 30, 32, 34, 40, 78, 183, 228
- Hanunó, people of the Philippines, 154, 155
- Harris, Roy, 179
- Hasan, Ruqaiya, 11, 25, 29, 36, 113, 147, 223
- Haviland, John B., 1, 65, 68, 90, 100, 153, 174, 181, 196
- Hawaiian kinship system, 157
- Hayashi, Makoto, 109
- Hazlehurst, B., 178, 188, 192
- Heath, Christian C., 178
- here-space, 33–66
- Heritage, John, 10, 92, 104, 106, 225
- Hesse, Friedrich W., 108, 149, 150, 179, 192
- heuristics for interpretation, 6, 15, 16, 17, 30, 31, 227
- Hewes, Gordon W., 68, 78, 86, 90
- Hill, Deborah, 28
- Himmelfmann, Nikolaus P., 25, 29, 30
- Hockett, Charles F., 27, 114
- Horn, Larry, 3, 31
- Hudson, Joyce, 73
- Hutchins, Edwin, 108, 149, 178, 181, 187, 188, 189, 191, 192
- Hutton, Christopher, 12, 13
- iconicity, 2, 13, 14, 16, 17, 18, 20, 26, 105, 106, 114, 115, 117, 140, 150, 155, 171, 190, 194
- implicature, 15, 16, 26, 27, 66, 106
- index-finger pointing, 38, 45, 51–5, 77–87, 85, 93, 106, 143, 151, 172, 198, 208, 209, 211, 215, 220
- indexicality, 2, 6, 12, 13, 16, 17, 18, 20, 26, 105, 114, 115, 117, 119, 123, 126, 145, 150, 151, 179, 185, 188, 209, 214, 219, 221, 222, 223, 225
- inference, 26, 30, 36, 37, 50, 65, 93, 179
- informational imperative, 92, 103, 106
- informative intention, 8, 14, 224, 226, 227
- informativeness scale, 31, 36, 50, 65
- insecure-reference utterances, 98, 99–107
- intersubjectivity, 110
- intonation, 1, 11, 108, 223
- intrinsic representations, 179
- isolating languages, 223
- Jackendoff, Ray, 2
- Jahai, people of Malaysia, 70, 72, 73
- Jakobson, Roman, 13, 15, 213
- Jefferson, Gail, 11, 99
- joint action, 27
- joint attention, 11, 13, 38, 39, 40, 44, 45, 49, 56, 58, 62, 64, 84, 90, 91, 97, 154, 169, 185
- Keenan, Edward L., 25, 31, 42, 58
- Kemmerer, David, 62
- Kemp, Jeremy H., 156
- Kendon, Adam, ix, 1, 9, 11, 16, 17, 18, 19, 33, 34, 48, 78, 88, 92, 124, 145, 153, 154, 184, 190, 225, 226, 228
- Key, Mary Ritchie, 72
- Keyes, Charles, 156
- Keysar, Boaz, 31
- kinship, 150, 153–77, 191, 193, 194–218
- Kirsner, Robert S., 29, 30
- Kita, Sotaro, xii, 1, 11, 12, 59, 68, 87, 90, 92, 150, 183, 190, 191
- Klima, Edward S., 115
- Kockelman, Paul, xii, 2, 3, 10, 12, 13, 17, 113, 227
- Kosslyn, Stephen M., 179
- Kuna, people of Panama, 68, 71, 74, 86, 87, 88
- Lakoff, George, 26, 28
- Lakoff, Robin, 29
- Lambrecht, Knud, 79, 93
- Langacker, Ronald W., 2, 28, 223
- Langston, W. E., 108, 150, 179
- Larkin, Jill H., 108, 149, 150, 178, 179, 192
- lateral axis, 156, 166, 168, 169, 172, 176, 183, 191, 194, 209
- Lawless, D. V., 154, 178, 184
- Layton, Robert, 149
- Lee, Penny, 27
- Levelt, Willem J. M., 114
- Levinson, Stephen C., xii, 2, 3, 12, 15, 16, 27, 28, 30, 31, 36, 50, 81, 84, 100, 104, 105, 106, 191
- Levy, Elena T., 97, 105, 106, 107, 123, 124, 145, 151, 184, 226

- Lewis, David K., 27, 30
 Liddell, Scott K., 1, 114, 118, 127, 146, 147,
 151, 152, 181, 184, 196, 203, 222
 Lillard, Angeline S., 5
 linear segmentation, 113, 114
 linearization, 19, 108, 113–15, 129, 143, 148,
 162–4, 179, 182, 183, 192
 linguistic typology, 66
 linguistics, 28, 108, 227, 228
 lip-pointing, 68–90
 Liszkowski, Ulf, 11, 29, 90, 91, 97, 98, 107,
 109
 location-focus utterances, 79, 93, 95, 97, 103,
 109
 lock-off, 224, 226, 227
 Lock-off, 226
 lock-on, 16, 224, 225
 Lock-on, 224
 Lucy, John, 191
 Luff, P., 178
 Lyons, John, 28, 31, 114

 Mandel, Mark, 18, 19
 marked signal, 11, 16, 32, 36, 76, 100, 116,
 158, 196, 200, 211, 225
 marriage, 155, 161, 173, 175, 176, 194, 195,
 205, 206
 Marshack, Alexander, 149
 Marslen-Wilson, William D., 151
 Mathur, Gaurav, 114, 118, 147
 maxim, 15, 30, 59, 66, 105
 Mayan, people of Central America, 191
 McNeill, David, 1, 2, 8, 10, 11, 14, 17, 106,
 107, 108, 113, 114, 116, 127, 146, 147,
 151, 152, 172, 184, 195, 198, 226, 227
 Mead, George H., 90, 92
 medial, distinction in demonstratives, 59
 Meira, Sérgio, 71
 Merleau-Ponty, Maurice, 194
 metaphor, 156, 194, 212, 219, 220
 metonymy, 120, 126, 129, 140, 145
 Miller, C., 115, 130
 Miller, George A., 145, 178
 Mitterer, Holger, 108
 Mopan, people of Belize, 191
 moves, 1, 8, 11–12, 20, 91, 105, 108, 113, 115,
 120, 129, 143, 146, 147, 148, 151, 184,
 188, 190, 215, 220, 221, 222, 227, 228
 Müller, Cornelia, 1, 18
 multimodality, 108
 Munn, Nancy D., 149, 154

 narrated event, 15, 213
 naturalness of semiotic mapping, 179
 Navajo, 72, 86, 87

 Neapolitans, 88
 Noble, W., 184
 non-conventional meaning, 6, 12–13, 12, 15,
 20, 93, 100, 114, 150, 221–3, 225, 226,
 228
 non-dominant hand, 115–18, 124–30, 135,
 137, 139, 140, 143–7
 Norman, Donald A., 108, 149, 178, 179, 181,
 184, 189, 190, 191
 norms, 10, 12
 Núñez, Rafael E., 191

 observer perspective, 128
 Ockham's principle of simplicity, 36
 Oestermeier, Uwe, 108, 149, 150, 179, 192
 Ogston, W. D., 1
 Okrent, Arika, ix, 7
 Olsher, D., 187
 Olson, David R., 149, 184
 Ong, Walter J., 149, 150
 onomatopoeic expressions, 26
 origo, 28, 58, 65, 93, 179, 183
 over-telling, 104, 105, 110
 Özyürek, Asli, 1, 9, 58, 59, 185

 Palen, L., 189
 Palmer, S. E., 179
 Parkin, Robert, 153
 Parmentier, Richard J., 2, 12
 Particularized Conversational Implicature, 15
 Pawley, Andrew, 11
 Peirce, Charles S., 2, 12, 13, 16, 17, 26, 27, 90,
 227
 perception, 108, 145, 178, 181, 219
 perimeter, of interactional engagement area,
 34–67
 phonological synchrony rule, 146
 phonology, 2, 26, 114, 115
 Pitjantjatjara, people of Central Australia, 73
 pointing, 11, 13, 16, 20, 29, 31, 38, 51, 52,
 62, 65, 67, 68–110, 135, 151, 152, 153,
 163–81, 165, 190, 198, 206–20, 222,
 223, 226
 Potter, Jack M., 156
 Povinelli, Daniel J., 90
 pragmatic synchrony rule for gesture, 146
 pragmatics, 3, 15, 25, 26, 27, 28, 31, 59, 66,
 91, 105, 109, 114, 117, 146
 Preissler, Melissa A., 5
 presupposition, 30, 79, 93, 95
 primitive, 11, 29, 35, 65, 90
 problem-solving, 91, 149, 179
 progressivity, 103, 109
 pronouns, 145, 156, 158
 prosody, 117, 123, 146, 147

- proximal, distinction in demonstratives, 25,
32, 42, 44, 50, 52, 59, 66, 196
- Quine, W. V. O., 90, 98, 151
- quotable gestures, 17
- rationality, 1, 15, 16, 223, 224, 225
- Raymond, Geoffrey, 92, 106
- recipient design, 184
- recognitionnal reference, 29, 30, 81, 83, 87, 88,
99, 101, 103, 104
- Reddy, M. J., 178
- register, 7, 87, 88
- Reilly, Judy S., 222
- relativity, 106, 191
- relevance, 15, 51, 227
- replica, 13
- representational determinism, 190, 191
- representational gestures, 58
- residue, 114, 147
- rest, of the hands as discourse signal, 141, 145
- Richert, Rebekah A., 5
- Rio, Knut, 154
- Robinson, Jeffrey D., 104, 109
- Rogers, Yvonne, 150
- Rolfé, Leonard H., 90
- Rosaldo, Michelle Z., 154
- Roszak, Theodore, 228
- Roth, Wolff-Michael, 154, 178, 184
- Sacks, Harvey, 9, 10, 11, 51, 54, 58, 59, 81,
92, 108, 178
- sagittal axis, 166, 168, 183, 202
- Sandler, Wendy, 115, 117, 123, 146, 147
- Sanford, Anthony J., 226
- Saussure, Ferdinand de, 10, 12
- Saussurean approach to meaning, 2, 113, 228
- Scaife, Mike, 90, 107, 150
- Scheffen, Albert E., 33, 42, 153
- Schegloff, Emanuel A., 1, 9–11, 51, 54, 58,
59, 81, 92, 103–4, 108, 178, 221, 225
- Schelling, Thomas C., 27, 30, 61
- Schutz, Alfred, 9
- Searle, John R., 10, 11
- semantic invariance, 13
- semiotic dimensions, 17, 18, 114
- semiotic unification, *see* unification of meaning
- semiotics, 2, 17, 113, 165, 228
- sentential mode of representation, 179
- shared intentionality, 11, 90
- Sherzer, Joel, 68, 71, 86, 87, 88, 90
- siblings, 157–77, 159, 160, 170, 172, 173, 174,
183, 191, 195, 198, 207, 210, 212–18
- sign, 2, 12, 13, 15, 17, 26, 27, 28, 67, 105, 107,
113, 227
- sign filtration, 16, 224, 225, 226
- sign language, 8, 114, 115, 129, 146, 147, 151,
152, 184, 222
- Silverstein, Michael, 2, 13
- Simon, Herbert A., 108, 149, 150, 178, 179,
192
- singularity, 13, 228
- skewing in kinship system, 157, 159, 160, 161
- Slama-Cazacu, Tatiana, ix, 1
- sociokinesics, 87
- speech event, 13, 15, 213
- Sperber, Dan, 3, 15, 153, 226, 227
- static view of meaning, 2, 10, 113, 147, 227
- Stivers, Tanya, xii, 32, 58, 81, 104, 109
- Stokoe, William C., 115
- Stolz, Christel, 183, 190, 191
- stop order, 226
- Streeck, Jürgen, 1
- Street, Brian V., 149
- Sturt, Patrick, 226
- Sun Tzu, 221
- surrogate space, 127, 152
- Sweetser, Eve, 191
- symbolic indexicals, 12, 13, 15, 19, 20, 67,
93, 95, 96, 150, 167, 168, 188, 222,
225, 226
- symmetrical gesture, 114, 116, 120, 121, 125,
128, 130, 131, 138, 139, 142, 143, 183
- symmetry condition, 115
- symmetry-dominance construction, 114–47,
147, 148
- syntax, ix, 108, 109, 113, 222
- Talbot, Henry Lee, 4, 5
- Talmy, Leonard, 18, 114
- Taub, Sarah F., 152
- Thai, 156, 158
- Thorpe, W. H., 114
- thumb pointing, 80, 93, 94
- timeless meaning, 13
- Tiriyó, people of Northwestern Amazonia, 71
- token space, 128, 152, 196
- Tomasello, Michael, x, 11, 27, 29, 56, 84, 90,
97, 109
- tracing, 19, 92, 124, 130, 131, 139, 151, 205,
206, 226
- Traweek, Sharon, 149
- try-marking, 108, 109
- Turkish, 59
- turn-constructional unit, 11
- type/token distinction, 13
- under-telling, 92, 104, 105, 110
- unification of meaning, 1, 6, 7, 16, 17, 25, 221,
223, 224, 227

Van Valin, Robert D., 11
Vermeerbergen, Myriam, 130
Versante, Laura, 88, 92
viewer perspective, 194
virtual reality, 20, 179, 183
Vygotsky, Lev S., 87

Walbiri, people of Central Australia, 154
Wassmann, Jürg, 149
Wierzbicka, Anna, 2, 13, 28, 29, 36, 184

Wilkins, David P., xii, 18, 19, 28, 32, 36, 68, 73,
78, 86, 88, 90, 92, 96, 149, 190, 226, 227
Wilson, Deirdre, 3, 15, 226, 227
Wittgenstein, Ludwig, 4, 23, 90, 98, 113, 151,
152
Wundt, Wilhelm, 1, 86, 90

Yucatec, people of Mexico, 191

Zhang, Jiajie, 108, 149, 150, 178, 179, 181, 190